JOURNAL *of* EDUCATIONAL RESEARCH

Volume XXVII SEPTEMBER, 1933 Number 1

THE VALIDITY AND RELIABILITY OF EDUCATIONAL DIAGNOSIS[1]

LEO J. BRUECKNER
University of Minnesota

My purpose in discussing the topic of educational diagnosis is to present some of the types of considerations that are to be the basis of a yearbook on that subject sponsored by *The National Society for the Study of Education*, and, in the name of the committee, to invite the active coöperation of those interested in this field of research. The subject has so many varied angles that it will be impossible to do more than merely sketch some of the problems that educational diagnosis faces if it is to be made valid and useful in locating the causes of inability to learn.

By diagnosis we mean the procedure that is used to identify or to determine the nature of a deficiency or weakness. The field of medicine presents the most outstanding contributions to the techniques of diagnosis. In medicine, emphasis has shifted from the mere observations of symptoms to the systematic application of techniques whose aim it is to determine the causes underlying these symptoms. Locating the cause or causes makes prognosis possible.

Just as the physician must have a well-classified knowledge of symptoms of various physical and mental ailments and, if possible, a knowledge of the causes underlying them as the basis of diagnosis, so the teacher should have a systematic knowledge of the most common pedagogical faults and their symptoms and causes in order to make a valid diagnosis of a pupil's inability to progress at a satisfactory rate. It seems that, in spite of as carefully guided instruction as we are able to give, there are pupils who for one reason or another fail to master the subject matter and skills that are being taught. Formerly it was our practice to fail such pupils. In recent years, definite attempts have been made to alleviate this condition by painstaking studies of learning

[1] Address delivered before the American Educational Research Association in Minneapolis, February 27, 1933.

difficulties and faulty reactions of pupils. As a result of these studies there has been assembled a very large amount of information helpful in educational diagnosis. It is the purpose of the yearbook to array this information in a systematic way and to make it available for workers in the field. At the same time, the committee will scrutinize the work that has been done, suggest its limitations and deficiencies, and point out the types of research that are needed to solve questions that are still unsolved, in many cases practically untouched.

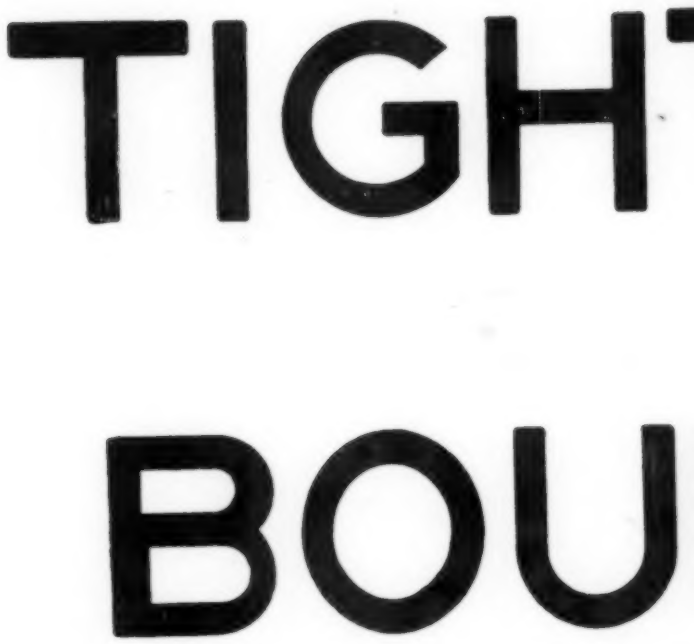
The need of such an array of information is obvious when one recognizes the fact that any diagnosis is conditioned by the previous training and experience of the examiner. Any diagnosis may be incorrect because of the inability or failure of the examiner to recognize the signs or symptoms of the defiicieney, or because of ignorance of fundamental facts concerning learning revealed by scientific inquiry; he may lack knowledge of the factors underlying the learning activity involved and ways in which they operate or are related; he may use techniques and devices that are inadequate, careless, and unscientific; the conditions under which the examination is given may be unfavorable or may be inadequately controlled, hence the responses the pupil makes may not be those that he typically makes in the ordinary activities in which they occur; the basis on which the diagnosis is made may not be sound, due to the fact that important items may be overlooked. Making available for teachers reliable objective techniques of diagnosis and training teachers in their use will do much to help them to eliminate or at least control those elements in the learning situation that are causing the difficulty.

It is interesting to note that some of the first attempts at the diagnosis of the causes of learning difficulty were made by physicians. The investigations of such men as Binet, who worked with subnormal children, come to mind. The removal of such obvious causes of difficulty in learning as faulty vision, poor hearing, and similar unfavorable physical conditions has often been possible. Medical men are obviously not equipped to locate the causes of many of the learning difficulties that arise in the school. Undoubtedly many of their findings may have important bearings on these problems which we at present cannot see. It is hoped that those interested in the solution of learning difficulties will make much more use of the techniques and findings of medicine, especially those of bio-chemistry, than they have in the past. It may be that there is a much more intimate relationship between physical maladjustment and learning difficulty than we have thought existed.

One of the very useful contributions to the diagnosis of learning difficulty has been the development of standard tests of intelligence. If one accepts the definition of intelligence as the ability to learn and assumes that these tests measure this aspect of intellect, one can readily determine the relative learning ability of an individual. If he scores low on these tests, one may say that at least one factor underlying his inability to learn may be lack of native ability. During the past few years there has raged a controversy between those who maintain that the mental level of an individual is determined by nature, while others insist that nurture, that is, environmental factors, are basic. If one assumes that low mentality as measured by mental tests automatically implies unmodifiable inability to learn, the task of teaching those of low mentality may be regarded as more or less hopeless. If one admits that our present tests of intelligence may, in many ways, be inadequate, possibly invalid, undue stress will not be placed on their results. Recent studies in the field of learning show that there is much more hope of teaching relatively subnormal children basic essential knowledge and skills than has often been thought possible. It is likewise true that often children of apparently superior intelligence encounter marked learning difficulties. It is thus clear that lack of intelligence alone cannot be regarded as the cause of inability to learn. Other factors are at work that must be determined.

To determine with precision the general level of a pupil's achievement in general and in particular subjects or phases of subjects standard tests of many kinds have been devised. By means of these tests it is possible to locate quite definitely the specific phases of particular subjects in which a pupil may be inferior. One difficulty in interpreting the results of many of the present day tests is the size of the units they use. Many of them are so large that a single unit represents the average growth which takes place in a whole semester. At present we have practically no achievement tests that measure in fine enough units progress from month to month, much less from week to week, or from day to day. That such tests are needed is obvious.

One of the findings of measurement by these achievement tests that has important bearing on diagnosis is the apparent specificity of the various learning activities and responses which constitute such general abilities as reading or arithmetic. Experiments have demonstrated that specific skills, such as rate of reading or ability to select the main topics



outcomes that may be more valuable. Certainly it is true that a diagnosis based on a mere achievement test of knowledge and skills is not a valid one, nor is it penetrating. Other outcomes must be considered.

The fact of variability of performance must be borne in mind in any interpretation that may be made of the results of most of our detailed diagnostic tests. A common procedure is the analysis of test papers by the teacher to determine the specific difficulties of pupils. For example, the teacher examines a test paper in arithmetic to determine the types of examples that were solved incorrectly. After this information has been secured, the teacher assigns practice exercises on similar types on the apparent assumption that failure to solve an example correctly is proof that the pupil does not know how to solve similar examples, or at least needs practice on them.

papers to discover the types of errors the various pupils made. Then he made a similar analysis of the errors found in their written work in all classes. He concluded that the results of the diagnostic tests did not furnish a valid basis of diagnosis since the pupils did not make the same kinds of mistakes in their usual written work that they made on the tests. He also found that there were errors on the regular work that did not appear on the test. To insure validity, a test of ability to punctuate should adequately sample the range of situations in which the skills involved in punctuation may be used. The test should also sample the varying reactions of the pupil to the same situation. For example, the pupil may be able in one case to punctuate today but not tomorrow. Some provision must be made for choosing sample situations which include the known variables which are most likely to affect the

of paragraphs, will readily respond to corrective measures or practice without apparently affecting in any way the other skills which go to make up the general ability to read. Profiles showing the comparative level of development of the various skills for individual pupils are highly irregular and uneven. This variability may be due in part to differences in training, possibly to innate differences in the individual himself, possibly to inaccuracy of the measuring instrument. The chief point to emphasize concerning this condition is the great variability in all human performance. This same phenomenon is also characteristic of the physical mechanism and all of its reactions. From the point of view of diagnosis this variability in performance is very significant. It suggests the inadequacy of any diagnosis based on a general test of performance. It seems essential to prepare a series of reliable tests of specific skills much more carefully constructed than any that are now available. The scope of these tests should be broadened so as to include the outcomes of learning such as interests, appreciations, ability to apply, and the like which to many are even more important than the outcomes that we are now able to measure. Tyler, for example, has proposed that instead of merely preparing tests to measure the information pupils may acquire, tests of other important objectives of a course should be devised. For example, under his directions there have been developed at Ohio State University tests for such outcomes of courses in biology as ability to apply the scientific method, ability to make inferences, ability to devise experiments, etc. This procedure tends to shift the emphasis in instruction from such relatively narrow objectives as knowledge of factual information, so often stressed in such courses, to those outcomes that may be more valuable. Certainly it is true that a diagnosis based on a mere achievement test of knowledge and skills is not a valid one, nor is it penetrating. Other outcomes must be considered.

The fact of variability of performance must be borne in mind in any interpretation that may be made of the results of most of our detailed diagnostic tests. A common procedure is the analysis of test papers by the teacher to determine the specific difficulties of pupils. For example, the teacher examines a test paper in arithmetic to determine the types of examples that were solved incorrectly. After this information has been secured, the teacher assigns practice exercises on similar types on the apparent assumption that failure to solve an example correctly is proof that the pupil does not know how to solve similar examples, or at least needs practice on them.

Recent studies we have made here at the University of Minnesota and elsewhere raise questions as to the validity of this procedure. An easy way to demonstrate unreliability of this type of diagnosis is to give the pupil more than one example of a given type to solve. In our experimental studies of this point we have given four examples of six different types in multiplication of fractions to solve. The examples were arranged in random order on the test paper. The pupils were given as much time as they needed to work the examples. No corrections were allowed. The papers were scored in such a way that we were able to record whether the pupil missed none, one, two, three, or four examples of each of the six types. It was found that in 1101 cases the pupils missed one or more of the four examples in each group of four. In 443 or 40.2 percent of the cases only one of the four examples was solved incorrectly. The other three were solved correctly. This suggests a large number of chance or accidental errors. In only 21.6 percent of the cases were all four of the examples of a single type solved incorrectly, thus indicating a difficulty that persisted and that was not due wholly to chance. Obviously therefore, it is not at all likely that a valid diagnosis of difficulty or inability in arithmetic can be made on the basis of a test containing only one example of a type. At least four examples should be used.

Another illustration of this factor of variability of performance of pupils is found in the study of diagnosis of difficulties in English by Willing. Willing gave a series of diagnostic tests in punctuation and other phases of English to a large group of pupils. He analyzed these papers to discover the types of errors the various pupils made. Then he made a similar analysis of the errors found in their written work in all classes. He concluded that the results of the diagnostic tests did not furnish a valid basis of diagnosis since the pupils did not make the same kinds of mistakes in their usual written work that they made on the tests. He also found that there were errors on the regular work that did not appear on the test. To insure validity, a test of ability to punctuate should adequately sample the range of situations in which the skills involved in punctuation may be used. The test should also sample the varying reactions of the pupil to the same situation. For example, the pupil may be able in one case to punctuate today but not tomorrow. Some provision must be made for choosing sample situations which include the known variables which are most likely to affect the

results. This applies to tests in all subjects. At the present time there are almost no tests constructed from this point of view.

To aid in the interpretation of the results of diagnostic tests numerous lists of the types of errors made by pupils in spelling, composition, arithmetic, and other subjects have been laboriously compiled. Knowledge of the types of errors pupils make gives some insight into the difficulties they encounter. In some cases it has been possible to devise improvements in instructional materials which have helped to avert the incidence of these faults. It should be pointed out, however, that relatively little information as to the exact causes of these difficulties is available. Furthermore, very little has been done to determine the seriousness of particular faults, their cruciality, which ones may be ignored, or the conditions of which they are symptoms. Practically nothing has been done to determine whether or not there are patterns of errors which would make it possible to classify learners into groups according to their difficulties much as the physician has learned to associate groups of symptoms with various illnesses. We have very few tests to locate with any degree of precision or exactness many of the types of faults we have learned to recognize in a vague way through observation of pupil responses. One may well wonder whether or not it will ever be possible to develop tests of causes of learning difficulty that will be as infallible as the well-known Shiek or Dick tests in medicine. The latter test fairly stable conditions, while learning is a growing, changing, cumulative condition.

It is interesting to note that very few of those who have made such painstaking studies of pupil errors have investigated the extent to which a particular fault detected in the pupil's work persists in his work. Unless an error persists it may be merely a chance one. The fact that a pupil misses four examples of a given type is no assurance that he misses them all for the same reason, although our studies show that in such cases there is usually present a constant error indicating a general difficulty. The interesting thing we find is that such constant errors are usually associated with other errors of a highly variable type. Apparently, pupils make a large number of random errors of a type that is difficult to explain in some cases, except on the grounds of sheer ignorance or inability to do the work. That this finding is probably true for other subjects than arithmetic was well illustrated in the findings of Willing, previously described. At present we are extending these studies of persistence of error into other fields. We are also study-

ing the effect of tests constructed in various ways on the persistence of errors by pupils. These studies have not yet progressed far enough to make results available.

At present we know little concerning the effect various kinds of motivation may have on the reduction or increase in particular kinds of errors by pupils. We know that positive forms of motivation greatly influence the rate of learning. It is conceivable that many of the types of errors that we find in the detailed lists would not be found at all if we could devise methods of controlling the efforts made by pupils on test exercises. Many of the errors are undoubtedly due to what we call mere carelessness or lack of interest. Under such conditions any diagnosis may be wholly invalid. Without the coöperation of the learner test results may give a picture that is wholly false. Much more needs to be done to standardize the conditions under which diagnostic studies are made, so that we may study more definitely the factors that influence pupil performance and contribute directly to the incidence of error.

The question may also be raised as to the extent to which errors and learning difficulties will disappear in the natural course of events, due to such factors as maturity, biological growth, routine teaching, and developing insight on the part of the learner. If faults disappear in the natural course of events without any diagnosis or special types of remedial attention problems of instruction are greatly simplified. If they persist, the need of diagnostic measures is obvious. It may be assumed that we shall greatly reduce learning difficulties as we perfect instructional procedures and materials. That we shall be able to eliminate such difficulties entirely is wholly unlikely due to our lack of knowledge of the various factors that condition learning. A simple technique we are using to study the relation of such factors as growth and routine teaching to persistence of errors in arithmetic is to analyze the work of pupils on identical tests repeated at varying intervals of time, such as an hour, a day, a week, two weeks, etc., to determine the extent to which difficulties disappear.

Bio-chemistry shows that the functioning of the internal secretion mechanism has a direct bearing on maladjustments of many kinds, both physical and mental. Since learning is attended by some modification in the nervous mechanism it seems likely that in many instances learning must be conditioned by the various glandular disturbances and imbalances that we know exist. For example, it is known that stuttering is often

associated with thyroid deficiency. At present we know very little concerning the relationship of glandular functioning and learning. This whole field merits extended research. For example, motivation which is recognized as one of the vital elements in learning is intimately related to emotion which is in turn controlled by glandular functioning. Many interesting hypotheses on the bearing of this fact on instruction could be proposed. For example, the variability of human performance may be intimately associated with chemical changes, for instance in the blood, or may be due to subtle forms of reaction to unknown forms of stimuli. If bio-chemists and those interested in learning problems could be brought together in the study of educational problems it seems as though there might result a whole series of findings of vital importance in diagnosis. This type of research would be especially helpful in getting at some of the causes of inability to learn. We have been too much concerned with the mere manifestations of difficulty through external symptoms. The time may be here when we shall shift to the definite attack on their causes. Much of our remedial work is ineffective since we know so little concerning causes.

If one scrutinizes the literature closely one can find here and there evidence that we are beginning to make definite attacks on the problem of discovering the causes of learning difficulty. We have long associated so-called poor eye-movements with reading disability. Often poor eye-movements have been thought to be the cause of poor reading. The remedial work to overcome this fault has been directed toward the improvement of eye-movement. Recent research suggests that certain types of faulty eye-movements should not be regarded as the cause but rather as a symptom of poor reading. Eames and Selzer have found that poor reading is definitely associated with a physical condition called muscular imbalance of the eye, technically termed "exophoria" or "heteropheria." Although compensation can clear vision, under these conditions an excessive amount of nervous energy may be required, leading to fatigue. As a consequence fluctuation occurs, which tends to increase fatigue. Faulty eye movements may be habits conditioned by this wavering, hence due to physical conditions, which would have to be corrected before correct reading eye movements could be developed.

Another interesting finding showing the relation of physical factors to learning is the apparently well established fact that lateral handedness or dominance seems to be associated with reversals in reading. Dear-

born reports that in some cases of reversals, if the pupil learns to read and write at the same time as is usually the case, the difficulty is aggravated. He reports the practice used in some schools of delaying the teaching of writing to first grade pupils who are left handed until they are well advanced in reading. This would tend to check the confusion of habits. To the casual observer, lateral dominance and disability in reading seem to be far removed one from the other, yet there may be here a subtle relationship of great significance in diagnosis.

Dearborn also reports that the tendency of poor readers to alter the order of letters in parts of a word, as reading "framing" for "farming" is associated with deviations from the usual conditions of lateral dominance, such as ambidexterity of hand and eye, change of handedness and eyedness, and the combination of right handedness and left eyedness or the reverse. It is not known whether there is here a causal relationship. If there is, it would seem almost futile to direct corrective efforts to work removing an apparent reading difficulty when it is not due to reading inability as such but to a physical condition or deficiency which conditions it. The errors due to reversals may be considered to be the consequence of conditions in the physical mechanism beyond the pupil's control.

Knight and Ford recently reported the results of an analysis of the errors found on particular multiplication combinations as affected by the depth to which particular combinations are embedded in an example. For instance, the combination 8 times 9 is embedded more deeply in the example 974 times 86 than in the example 749 times 68. In one case the combination is the last one while in the other it is the first one. Knight and Ford found that the more deeply the combination was embedded in the example the more likely pupils were to make an error in working it. Thus a pupil may know a combination in isolation but fail to give it correctly in a long example. The authors believe that "This difficulty may be thought of as a temporary lapse in, or degeneration of, ability within a segment of work. Thus in addition to such factors as subject matter, age, motive, sex, previous fatigue, fluctuations in health, and the like much of error may be the direct and protective outcome of a disregard on the part of the pupil or teacher for the periodicity of output of energy during the working period, even if it may be of only a short duration." The cause of errors cannot be fatigue since in succeeding examples the errors on the same combinations do not reoccur.

Many other illustrations might be presented of factors that suggest the difficulty there may be in making a valid diagnosis of causes on inability of pupils to learn or of the errors they make. As Knight and Ford say, "We badly need an adequate theory of the nature of error. It is apparent that much of the present thinking about error assumes that error is a much simpler phenomenon than it is. Errors may not be a single thing in their nature. Treating all errors alike is analogous to treating all fevers alike—a purely makeshift and palliative futility. Thus, the therapy error due to temporary lapse is not the therapy for actual ignorance. Much of our treatment of error, be it ever so heroic, is nevertheless superficial and unsuccessful because it is based on too little understanding of the psychological nature of the error and the processes leading to it."

In the forthcoming yearbook it is proposed to study the problems of diagnosis from many angles. This can readily be observed by an analysis of the variety of interests represented on the committee. The following are the personnel: G. T. Buswell, University of Chicago; Ralph Tyler, Ohio State University; Willard Olson, University of Michigan; Lee Travis, State University of Iowa; P. T. Rankin, Detroit, Michigan; J. L. Stenquist, Baltimore, Maryland; and Leo J. Brueckner, University of Minnesota, *Chairman*. The committee plans to ask the coöperation of specialists in all fields in the preparation of what it hopes will be an authoritative statement of the present status of diagnosis, practical methods of applying valid diagnostic techniques, the difficulties and limitations of diagnosis of learning problems, and needed research. The writer suggests that anyone who is at work on some piece of research that seems to bear on any of the problems herein outlined or others of a similar nature get in touch with the members of the committee who are especially interested in the particular field. It is the desire of the committee to make this yearbook, in a real sense, a coöperative undertaking.

AN INTEREST INVENTORY FOR HIGH SCHOOL PERSONNEL WORK

GERTRUDE HILDRETH
Lincoln School of Teachers College
New York City

THE popularity of the inventory as a means of studying personality traits is indicated by the increasing frequency with which this technique is employed in personnel work. Casual observation is recognized as inadequate for the study of adolescent problems. Successful adjustment of the pupil to the school situation requires that teachers and advisors have adequate information concerning the child's interests and preferences with respect to a large number of activities and fields of interest, companions, sports, social activities, reading, study habits and school subjects and vocational preferences. Knowledge of the child's habitual modes of thinking and acting as well as his own evaluation of himself are essential to the full understanding of the child and his problems. Without such information, teaching and guidance work may be not only ineffectual but actually harmful in its consequences.

ALTERNATIVE METHODS OF STUDYING PUPIL INTERESTS

There are several alternative methods of obtaining pupil interest and attitude data. One method is that of personal interview with the pupil. Another is observation of the individual while he is engaged in a variety of activities in and out of school. A third method is that of asking the pupil to record on a suitably arranged blank statements concerning his activities and preferences. If there are many cases to study or if information for a large proportion of pupils is required, the first method, used exclusively, is almost out of the question unless the school has a large staff for guidance work. Aside from the amount of time consumed there are the additional problems of pupil reticence during interviews and the recording of results in objective form, uncolored by the preferences and biases of the interviewer.

The second method is even more time consuming and just as dependent upon highly trained observers for reliable work. It avoids the factor of embarrassment to the pupil, provided he does not discover that he is being observed. Either of the two methods represents a highly specialized activity. The third method, that of having pupils record in

some objective fashion information concerning their personality traits, interests and habits, though not in all respects free from objectionable features from the standpoint of reliability and validity, nevertheless is a time saver, gives the pupil a part in the undertaking, and affords permanent records in convenient form for filing. When time and adequate personnel staff are available such records may profitably be supplemented by data obtained through methods one and two. A record blank filled out by the pupil constitutes a natural starting point for an interview since it releases tension on the part of the child by centering attention on his paper rather than on himself. The complete "blocking" which even the expert interviewer encounters when he puts such questions as "Now tell me which subjects you have most difficulty with" is in decided contrast to the situation when the interviewer can say "Now, let's see what you wrote down about your school work. Here you have the names of the subjects you like best, and here the ones you do not like so well." At this point the pupil will usually volunteer more information about his success in school work.

The objection to the questionnaire method when the record is made by the child and is unchecked by other methods is that the child may be unable to give an accurate picture of himself with respect to the traits in question or he may intentionally falsify his record to make it fit the picture he wants his teachers and advisors to have of him. Other things being equal, the pupil who is older and more mature mentally is more capable of giving an accurate picture of his personality. Whether, even though he is more capable of giving an accurate picture, he is more apt to do so, depends upon several factors: the morale of the school and the sportsmanship of the pupils, the use that pupils feel is to be made of their records, the purposes for which the information is being collected, the pupils' attitude toward the person who is directing the study, and the extent to which pupils are momentarily influenced by companions or "heroes" among their classmates. Results will be more satisfactory if pupils are convinced that they are to be helped in their problems rather than to feel that someone is trying to "get the goods" on them. The request that pupils fill out interest inventories should come from a person with whom the pupils are in satisfactory rapport.

DESCRIPTION OF THE INVENTORY RECORD BLANK

The inventory blanks to be described here have been used with satisfactory results at the Lincoln School of Teachers College, Columbia Uni-

versity, in obtaining information from pupils of high-school age for personnel work. The description will be confined to the use of the test blank itself rather than to the contribution to adolescent psychology of the facts discovered. For this latter interesting and pertinent topic separate treatment is necessary.*

The material given to the pupil consists of two parts: a single sheet on which he is to record his interests and preferences in a number of categories and a check-list which contains an extensive list of headings under each category. Only a few sample items in each category of the check-list are given here. The two blanks are given to the pupil with the statement that the school wishes to become better acquainted with him, and in order to do so would like to have him tell about some of his activities and interests. Directions for filling out the blank are then given. Copies of these blanks are shown below.

If the examiner wishes, the accompanying List of Activities and Interests need not be used. Instead, pupils are simply asked to write down the names of their preferred activities, reading material and the like, with only a general suggestion as to the type of information required in each category. One objection that has been raised to a check list of items is that pupils may record as their choice items that might otherwise never have occurred to them. On the other hand, since recognition is easier, psychologically, than recall, the use of the check sheet facilitates response. With no assistance to his memory the pupil's response becomes impractically slow, and he may be totally unable to recall his interest in a particular category, as for example the name of the magazine he most enjoys reading.

To the contrary, it may be argued that if the child cannot recall the item, then it cannot represent a real or abiding interest or preference. Research is clearly needed at this point to determine which method or combination of methods is most practical and at the same time most trustworthy. In the case of younger children there is danger of fatigue if pupils are requested to read through the long lists of items before recording their responses. In such cases the teacher or investigator might read quickly sample items chosen at random from the list under each category, and then ask pupils to list their preferences.

---

* See Hildreth, Gertrude. "Adolescent Interests and Abilities," *Journal of Genetic Psychology*, June, 1933.

PUPIL'S INFORMATION BLANK

Name..... Age..... Grade..... Date..... Boy..... Girl.....

Home Address..... School Last Attended.....

I. ACTIVITIES

Those liked	Those disliked
1.	1. .....
2.	2. .....
3.	3. .....

II. GAMES—SPORTS

Those liked	Those disliked
1.	1. .....
2.	2. .....
3.	3. .....

III. SCHOOL SUBJECTS (First, list here by number all subjects in the list that you have studied during the past two years).....

.....

Those liked	Those disliked
1.	1. .....
2.	2. .....
3.	3. .....

IV. TYPES OF BOOKS

Those liked	Those disliked
1.	1. .....
2.	2. .....
3.	3. .....

V. BOOKS READ DURING PAST TWO YEARS

1. .....
2. .....
3. .....
4. .....

VI. MAGAZINES PREFERRED

1. .....
2. .....
3. .....

VII. NEWSPAPERS READ

1. .....
2. .....

VIII. OCCUPATIONS YOU MIGHT LIKE

1. .....
2. .....
3. .....

IX. PEOPLE YOU LIKE TO BE WITH IN YOUR FREE TIME

Those liked	Those disliked
1.	1. .....
2.	2. .....
3.	3. .....

X. SCHOOL PREFERENCE AFTER HIGH SCHOOL.....

.....

XI. PERSONAL CHARACTERISTICS

Those that describe you	Those that do not describe you
1.	1. .....
2.	2. .....
3.	3. .....
4.	4. .....
5.	5. .....
6.	6. .....
7.	7. .....
8.	8. .....
9.	9. .....
10.	10. .....

XII. If you could have your wish for the thing you desire most what would it be?

.....

.....

.....

LIST OF ACTIVITIES AND INTERESTS

(Referred to as the check list)

Directions: DO NOT WRITE ANYTHING ON THIS BOOKLET. This list is to be used as an aid in filling out your INFORMATION BLANK. Read over the list of items in the LIST OF ACTIVITIES and INTERESTS, decide on your first choice, write its name and number on the first line of your INFORMATION BLANK under the topic Activities. Then write the number of your second choice. Do not hesitate to write down on your INFORMATION BLANK any information you wish to give even though it may not be given in the typed list. For example, if you enjoy reading most a magazine not mentioned here, be sure to list that magazine as your first choice. Do not feel that you must give information simply because the topic is included in the blank. If you are not going to school after high school, you would of course leave out information on this item. Do not spend too much time thinking about any item or topic.

I. ACTIVITIES

1. Watching athletic events
2. Looking up information
3. Taking care of children
7. Playing piano
16. Working around machinery
22. Driving an automobile
25. Inventing things
26. Visiting museums
30. Attending movies
34. Bicycle riding
37. Going to parties
49. Studying
51. Hating "dates"
52. House work
61. Canoeing
64. Reading
69. Swimming
72. Woodcraft

II. GAMES—SPORTS

1. Anagrams
5. Baseball
6. Basketball
10. Bridge
14. Crap Games
18. Football
21. Hockey
23. Ping Pong
33. Tennis
34. Track
37. Wrestling

III. SCHOOL SUBJECTS

1. Algebra
2. American History
8. Chemistry
9. Civics
10. Commercial Subjects
14. English
17. French
18. General Science
23. Household Arts, Cooking, etc.
26. Latin
29. Music—Singing
32. Physical Education
36. Shop Work
43. Zoology

IV. TYPES OF BOOKS

2. Encyclopedias and Books of Knowledge
3. Stories of home and school life
4. Nature and animal stories
5. Novels and love stories
9. Adventure stories
11. Mystery stories
15. Classics
16. History
17. Science
19. Sports
21. Humor
23. Arts

V. BOOKS READ DURING PAST TWO YEARS

List them on INFORMATION BLANK

VI. MAGAZINES PREFERRED

1. American
5. Century
9. Harpers
10. Judge
11. Liberty
12. Life
17. Outlook
20. Red Book
26. American Boy
29. Boy Mechanic
32. Cosmopolitan
36. Girl Scouts
37. Golden Book
40. New Yorker
49. True Stories
52. American Girl
56. Atlantic Monthly
58. College Humor
65. Literary Digest
78. Detective Stories
79. Good Housekeeping
82. Ladies Home Journal
91. Saturday Evening Post
97. Woman's Home Companion
98. Youth's Companion

VII. NEWSPAPERS READ

List them on the INFORMATION BLANK

VIII. OCCUPATIONS YOU MIGHT LIKE

1. Accountant
3. Actor or Actress
5. Architect
7. Artist
8. Athletic Director
9. Author
11. Automobile Repairman
13. Aviator
14. Baker
15. Banker
26. Carpenter
28. Chauffeur
29. Chemist
30. Civil Engineer
32. College Professor
40. Dentist
41. Detective
43. Doctor

44. Dressmaker
45. Editor
49. Factory Worker
55. Football Coach
58. Governess
60. Hairdresser
62. Homemaker
69. Janitor
70. Jazz Orchestra Player
74. Lawyer or Judge
75. Librarian
76. Life Guard
86. Merchant
87. Milliner
88. Miner
91. Motion Picture Director
92. Musician
95. Nurse
96. Office Clerk
101. Photographer
102. Playground Director
104. Poet
105. Policeman
106. Politician
107. Preacher or Priest
108. Plumber
110. Private Secretary
116. Real Estate Salesman
117. Reporter
121. School Teacher
122. Scientific Research Worker
127. Social Worker
128. Soldier or Sailor
131. Stenographer or Typist
136. Tailor
142. Undertaker
143. Waiter or Waitress

IX. PEOPLE YOU LIKE TO BE WITH IN YOUR FREE TIME

1. Older boys
2. Younger boys
3. Older girls
4. Younger girls
5. Boys your own age
6. Girls your own age
7. Mother
8. Father
9. Other adults
10. No companions
11. Little children
12. Teachers

X. SCHOOL PREFERENCE AFTER HIGH SCHOOL

1. Kindergarten Training School
2. School of Commerce or Business School
3. Music School
4. Boarding School
5. Dramatic School
6. Large Men's College or University
7. Large Women's College or University
8. Medical School
9. Law School
10. Dental School
11. Junior College
12. Aviation School
13. State University
14. State Agricultural School
15. State Engineering School
16. Normal School
17. Teachers College
18. Art School
19. School of Physical Education
20. Theological Seminary
21. Trade School
22. Large Coeducational College
23. Small Coeducational College
24. Small Women's College
25. Small Men's College

XI. SELECT THE PERSONAL CHARACTERISTICS THAT DESCRIBE YOU

1. Happy most of the time
4. Quick tempered
6. Like quiet activities
10. Afraid of many things
11. Dislike parties
14. Interested in people and their affairs
15. Able to work well with others
18. Considered a good sport
21. Have good taste in dress, decorations
25. Have skill in art, music, writing
26. Follow directions easily
30. Broadminded
34. Good at mechanical things
37. Slow but sure thinker
41. Usually on time
43. Hate to work
46. Plan ahead
49. Dislike to put things away when through
50. Fond of school work and activities
53. Have difficulty with school teachers
55. Considered a leader in school activities
56. Sleep well
58. Have many headaches
61. Cannot eat many vegetables
64. Considered a good dancer or skater
65. Enjoy being away from home
67. Have understanding and sympathetic parents
69. Dislike to do what parents ask
70. Have a better than average home

USE OF THE INVENTORY FOR GROUP OR INDIVIDUAL EXAMINATION

The inventory is equally suitable for either group or individual examining. In group examining the administrator should arrange the

situation so that pupils cannot communicate with each other while filling out the blanks. On one occasion a boy remarked to his companions, in a stage whisper, that in the list of vocations the occupation of "husband" and "father" was missing. The result was, in that particular group of papers, decidedly more listing of "husband" or "father" as the preferred occupation than chance would allow. In order to save time the examiner may request that instead of pupils writing out in full the names of their preferences, they indicate them by the numbers placed before the items in the check list. Thirty minutes is usually an adequate time allowance.

In individual examining, when the material is to be used in conjunction with other tests, our procedure is to give intelligence or achievement tests first, to follow these with the inventory described here, and in turn to follow the inventory with an interview starting from the information furnished by the inventory. Complete case work which will not be described here requires, necessarily, much further investigation than the outline of pupil study given here, but all of the information obtained by these methods may properly be considered part of a case study.

When the questionnaire follows standard tests the pupil is inclined to accept it impersonally as just "so much more" of the same thing. As he proceeds and finds that it is unlike the test he has just been having, he continues, often with considerable pride and gusto to furnish information about himself. He does not realize that later he will be interviewed concerning its contents, nor is he conscious of the fact that he is often "giving himself away" by his responses.

PRACTICAL USES OF INFORMATION FURNISHED BY THE INVENTORY

The advantages to the school in having more complete information of individual pupils such as the inventory described here provides are manifold. More pertinent advice can be given concerning the selection of courses, vocational guidance, choice of college, the arrangement of sports, and social and activity programs. The data are invaluable in talks with parents, in making individual assignments, in classifying pupils, in coöperating with health services, in arranging the pupil's study program, and in case study of problem pupils.

In order to interpret the responses of any one individual pupil it is highly desirable to have at hand data for a large unselected sample of the population of pupils of the same age or grade. Using these

"norms" as reference points, individual idiosyncrasies or abnormalities may readily be determined.

REFERENCE LIST OF INTEREST QUESTIONNAIRES AND PERSONALITY RATING SCALES

There are many variations of the interview and questionnaire method of determining interests and habits of high school and college students and workers in industry. For the benefit of the reader who wishes more information about such materials some of the most extensively used and practical devices for work with high-school pupils are listed here. Other adaptations are possible which these methods suggest.

Interest Analysis—Form A. New York City: Association Press, 1931.

Brainard, P. P. *Specific Interest Inventory For Boys—For Girls.* New York City: Psychological Corporation, 1932.

Garretson, O. K., and Symonds, P. M. *Interest Questionnaire for High School Students.* New York City: Bureau of Publications, Teachers College, Columbia University, 1930-31.

Freyd, M. *Occupational Interest Blanks.* Chicago: C. H. Stoelting and Company, 1922.

Hepner, H. W. *Vocational Interest Quotient Booklets.* New York City: Psychological Corporation, 1931.

Jenson, D. W. *Interest Questionnaires:* Genetic Studies of Genius, Volume III. Stanford University, California: Stanford University Press, 1930.

Kelley, T. L. *Interest Tests:* "Educational Guidance—an Experimental Study in the Analysis and Prediction of Ability of High School Boys." Contributions to Education, No. 71. New York City: Bureau of Publications, Teachers College, Columbia University, 1914.

King, L. H. *Personnel Record and Interest Test:* "Mental and Interest Tests." Contributions to Education, No. 444. New York City: Bureau of Publications, Teachers College, Columbia University, 1931.

Miner, J. B. *Analysis of Work Interest Blank.* Chicago: C. H. Stoelting and Company.

Minnesota Interests Questionnaire. Minneapolis: University of Minnesota Press, 1930.

Strong, E. K. *Vocational Interest Blank.* Stanford University, California: Stanford University Press, 1930.

Wyman, J. B. "Tests of Intellectual, Social and Activity Interests." *Genetic Studies of Genius, Volume I.* Stanford University, California: Stanford University Press, 1925, pages 455-483.

The most widely used personality trait inventories are the following:

Bernreuter, R. G. *The Personality Inventory.* Stanford University, California: Stanford University Press, 1931.

Thurstone, L. L., and T. G. *Personality Schedule.* Chicago: University of Chicago Press, 1928.

Woodworth, R. S., and Mathews, E. *Questionnaire.* Chicago: C. H. Stoelting and Company, 1923.

IMPROVING THE MEASURING QUALITIES OF AN ESSAY EXAMINATION

VERNER MARTIN SIMS
University of Alabama

THE PROBLEM

THE great variation found in the works assigned by different supposedly competent readers to the traditional or essay examination may be due to: (1) disagreement as to what is being measured, disagreement as to whether one shall consider content, organization, use of English, neatness, etc.; (2) disagreement as to the contribution of the various elements to the total, disagreement as to the weighting of the questions and the elements within a question as well as to how much such factors as those mentioned above should contribute; (3) differences in the standards of grading, differences among systems and among teachers using the same system; and (4) the subjectivity of the materials being marked. The first cause of variation is not inherent in the nature of the instrument but is the result of disagreement as to the ends of instruction or as to the function of the particular examination; the second is a common one in human measurement, where our units are never equal, but it is subject to certain control, if by nothing more than arbitrary agreement as to the weighting; and the third, not a matter of measurement but of the interpretation of the measurement, is due to an attempt on the part of the reader to express performance in terms of some real or imaginary, and wholly arbitrary, grading system. Theoretically, then, these three factors may be considered as extraneous sources of error, and the ultimate value of the essay examination can be determined only after we have developed means of controlling them or have demonstrated that they are beyond control.

Kelly (4*, p. 83) and Ruch and Fauber (7, pp. 102-106) have shown that the use of certain scoring rules, which would tend to reduce disagreement as to what is being measured and as to the contribution of the various elements to the total, will greatly reduce the variations among readers. Ruch (7, p. 105) says, "... the variability ... may be cut at least in half by the use of carefully laid-down scoring rules." Sims (12)

*Refers to list of references at the end of this article.

found that by considering the reader's marks as scores (raw units of measurement) and converting them to grades (standard units in the sense of expressing position in a group) the variability from reader to reader was reduced more than one-half.

Will the combination of these two procedures still further improve the measurement obtained by the essay examination? It is the purpose of this paper to evaluate an essay examination which was scored with rules and the marks then transmuted to grades.

THE EXAMINATION

An hour essay examination, consisting of ten questions based upon a unit in General Psychology, was given to a class of 33 students who had just completed the unit of work. The examination, a *bona fide* one, was to be combined with other measures to determine the preliminary grade for the semester. In an earlier study Sims (11) has pointed out that the questions commonly used in essay examinations may be classified as objective (simple-recall or short-answer) and discussion (criticize, discuss, explain, interpret, trace, etc.) types, and that about 60 percent of the questions are of the objective type. This examination was a typical one in the sense that it was made up of six objective and four discussion questions. The following questions are typical of the two types: (*objective*) "What are the chief divisions of the psychological field as given by Dashiell? Identify the nature of each with a sentence. Mention the name of one man prominently associated with each field." (*discussion*) "Discuss the significance of psychology for three occupational groups."

The students were assigned numbers to be used in identifying the papers, and were instructed to answer each question on a separate sheet of paper. They were cautioned to so distribute their time as to be able to answer all questions in the allotted time. All students managed to complete the examination, although a few of them were rushed at the end.

SCORING AND GRADING

Eight readers marked the set of 33 papers. The readers were three instructors, three 'readers,' and two students, in the Department of Psychology, College of Education, University of Alabama. The regular testing program in General Psychology consists of a series of coöperatively prepared objective examinations, so that none of the readers had had special practice in marking essay examinations in this subject. The instructors were, of course, familiar with the marking of essay examinations and two of the 'readers' had been reading essay papers in Educa-

tional Psychology. The other readers had had practically no experience with them. In scoring the papers all readers used the same scoring rules, and each reader's marks were afterward converted to grades by a uniform procedure.

The scoring rules involved general instructions to consider content only in the objective questions and content and organization in the discussion questions and then specific directions were provided for scoring the two types of questions. For the objective questions a key was furnished which gave the correct answer for simple-recall and acceptable answers for short-answer questions (the text being the authority), and arbitrarily assigned score values of one point for simple-recall items and two points for entirely satisfactory or one point for partially satisfactory short-answer items. A modification of a rating technique which the writer has described and evaluated in a previous paper (10) was used for scoring the discussion questions. Each question was scored by sorting the 33 answers on the basis of merit into five groups, a normal distribution being assumed and conformed to rather rigidly. The instructions given the readers were:

1. Read the section or sections in the text devoted to the question. (References were furnished.)
2. Quickly read through the answer given to the question by every student.
3. Go back and sort the answers on the basis of merit into five groups as follows: (conform rather closely to the numbers suggested for each group) (a) the 3 or 4 very best papers, (b) the 6 or 7 superior papers, (c) the 12 to 15 average papers, (d) the 6 or 7 inferior papers, (e) the 3 or 4 very worst papers.
4. Give a score of 10, 7, 5, 3, and 0 points, respectively, to the various groups.

The readers reported the score on every question for each paper. To determine the total score for a paper the questions were considered of equal value and were each assigned a score value of ten points. The scoring method used for the discussion questions automatically weighted them equally and gave this score value. Since the number of items in the several objective questions varied, it was necessary to express the score made as a fractional part of ten; which was done simply by dividing the score made by the possible score for the question and multiplying by ten.

Each reader's scores for the set of papers were converted to grades, with 70 percent as a passing grade, by a method proposed elsewhere by

this writer (9). This procedure involved locating the passing score at —1½ standard deviations and then graphing the relationship between scores and grades, assuming that the passing score corresponds to the passing grade and the maximum possible score corresponds to the maximum possible grade. The steps in the process were:

1. Locate the lowest passing score. (—1½ standard deviations.)
2. With the lowest passing score as the origin, lay off in equal steps on the vertical axis of a graph the possible scores.
3. With the lowest passing grade as the origin, lay off on the horizontal axis of this graph distances corresponding to the possible grades.
4. Draw a diagonal through the origin and the point corresponding to the maximum possible score and the maximum possible grade.

To find the grade of a particular score: Read across to the diagonal and up or down to the percentage line.

In order to evaluate the examination when so scored and graded we determined: (1) the variation from reader to reader in the grades assigned the respective papers; (2) the objectivity as shown by (a) the correlation between each reader and the average of all readers, and (b) the correlation between each reader and every other reader; (3) the reliability; and (4) the correlations between this examination and a new-type test covering the same subject matter. Evidence on the last three points was secured not only for the total examination but also for the two types of questions, objective and discussion.

VARIABILITY OF THE GRADES

The great variability which is usually found when a number of readers mark the same papers needs no comment here. To what extent has our treatment reduced this variability? Table I shows the grades given by each of the eight readers, the range from the highest to the lowest mark given, and the probable error, .6745 standard deviations (3, p. 211), for each of the 33 papers.[1]

The average range for the 33 papers was 4.4 points, the maximum range for any paper was 7 points, three papers having this range and four of the papers having a range of but 2 points. The average probable error for the 33 papers was 1.0 points, the maximum probable error being

[1] The first three readers in this table (and in other tables reporting the several readers) were the instructors, but there was no consistent difference between the marks of the instructors and the other readers so no further attention will be given to them.

TABLE I

SHOWING GRADES GIVEN BY 8 READERS TO A SET OF 33 EXAMINATION PAPERS WHEN SCORED WITH RULES AND THEN CONVERTED TO GRADES

Paper	Reader								Range	P. E.
	1	2	3	4	5	6	7	8		
1	94	96	93	94	94	93	94	92	3	.7
2	92	90	87	86	88	87	85	86	7	1.6
3	87	88	89	86	84	90	89	88	6	1.3
4	86	85	85	87	89	88	89	88	5	1.3
5	90	87	86	87	87	87	86	87	4	.6
6	85	86	85	86	85	86	84	88	4	1.0
7	85	86	85	85	86	85	84	86	2	.4
8	85	84	83	83	87	84	84	87	2	1.0
9	85	84	83	83	83	85	84	84	2	.7
10	83	83	86	82	82	83	85	83	4	1.1
11	83	83	82	82	82	82	81	86	5	.8
12	83	83	85	83	82	81	82	82	4	.8
13	82	82	80	81	82	86	84	81	6	1.3
14	82	82	84	82	82	83	80	83	4	.7
15	83	82	81	78	81	83	83	83	5	1.2
16	79	81	78	82	79	85	80	82	7	1.5
17	79	82	82	80	81	80	79	80	3	.9
18	83	79	82	79	81	82	80	76	7	1.5
19	81	81	78	80	78	83	80	80	5	1.1
20	79	82	78	79	79	81	80	79	4	.9
21	79	82	79	79	78	80	79	79	4	.7
22	79	80	76	77	78	76	77	78	4	.9
23	77	78	75	77	78	77	78	81	6	.6
24	81	77	80	74	78	78	76	75	7	1.6
25	75	76	73	77	73	73	76	76	4	1.2
26	75	76	72	75	74	75	73	75	4	.9
27	75	74	71	72	74	74	74	72	4	1.0
28	73	73	69	73	73	74	75	74	6	.9
29	71	70	71	74	70	73	70	73	4	1.2
30	69	71	69	69	71	69	72	70	3	.9
31	69	69	69	71	69	71	71	69	2	.9
32	69	69	68	72	69	69	68	71	4	.8
33	65	65	65	68	66	66	65	68	3	.9
Average	80.2	80.1	79.1	79.5	79.5	80.3	79.6	80.1	4.4	1.0

1.6 and the minimum .4. Interpreted in terms of deviation from the average, only five (1.9 percent) of the 264 grades reported in the table deviate from the average grade for the paper as much as 4 points, 95 percent of them deviate less than 3 points, 80 percent of them deviate less than 2 points, and 50 percent of them deviate less than 1 point.

As compared with other reported experiments on grading essay papers, the lack of variability here is so pronounced that it needs no particular comment. For example, a direct comparison between the spread here

and that found in the study already mentioned (12) seems valid. In this earlier study a set of twenty examination papers in general psychology were marked under the same conditions by readers with practically the same experience but with the difference that no instructions other than to mark on a percentage basis were given. The readers (four of them were the same as those used in the present experiment) were three instructors, two readers, and four students in the department, none of them with any specific training in reading papers in general psychology. The average range for nine readers marking with no instructions was 28.3 points, the maximum range for any paper was 50 points and the minimum range 10 points, and the average probable error was 5.9 points.

Assuming equally competent readers on these two examinations, one scored with rules and the other without, a comparison of the variability of the marks before and after conversion to grades makes possible an evaluation of the relative role in reducing variability played by the two procedures, scoring rules and conversion to grades. When the papers were marked with *no rules* and the raw scores considered, the average range was 28.3 points and the average probable error was 5.9; when converted to grades (by the method outlined above) the average range became 10.0 and the average probable error 2.2. When the papers were marked *with rules* and the raw scores considered, the average range was 13.1 and the average probable error was 2.8 points; when converted to grades, the average range became 4.4 points and the average probable error 1.0 points. The reduction in variability due to conversion to grades seems to be slightly more than that resulting from the use of scoring rules, but the significant fact is they are obviously independent sources of error since their combination reduces the variability much more than either alone reduces it.

OBJECTIVITY

Used in the sense of reliability of scoring, objectivity is determined by keeping the student performance (the papers) constant and varying the readers. This we have done, and the *lack of variability* reported in the previous section means the method of scoring and grading used for this examination gave rather objective measurement. To get quantitative measures of the objectivity two methods were used, correlation between a reader and the average of the other readers, and correlation between the several readers paired. Since the average of many readings gives a more reliable measure than any one reading, the former method would be expected to yield higher coefficients and to show less variability than the

latter, but the latter would probably more nearly duplicate the usual marking situation.

Table II presents the coefficients of correlation between the marks given by each reader and the average of the other readers, and the average coefficients for the marks given by each reader paired with those given by every other reader.

TABLE II

CORRELATION BETWEEN EACH READER AND THE AVERAGE AND THE AVERAGE CORRELATION BETWEEN PAIRED READERS FOR THE TOTAL EXAMINATION

	Reader								
	1	2	3	4	5	6	7	8	Average
Reader . average	.97	.98	.97	.96	.98	.96	.96	.96	.97
Average Reader . reader	.94	.95	.94	.94	.94	.94	.94	.93	.94

The coefficients between the marks of a reader and the average are consistently high, higher than those usually found. For example, Shriner (8) reports an average r of .94 (range from .87 to .97) between each reader and the average of 25 readings of a set of algebra papers, and an average r of .91 (range .82 to .96) between each and the average of 25 readings of a set of English papers. For the nine readings of the psychology papers, scored without rules, mentioned in the previous section (12), the average r between each and the average was .91 (range .87 to .97). The coefficients between the marks of paired readers are lower but the average for each reader is agreeably high. The distribution of the 28 coefficients (the possible number of combinations) from which these averages were determined ranged from .91 to .97 with an average of .945.

These correlations seem to indicate an extremely high degree of objectivity for the total examination. What is the relative objectivity of the two types of questions, objective and discussion? Because the method of scoring used means that the number of questions affected the objectivity, coefficients such as those determined above would not be comparable for the two types. In order to make them comparable two of the objective questions were eliminated at random, thus making the two types of equal length, and the correlation between the marks of each reader and the average of the eight for the two types of questions was determined. These coefficients are presented in Table III.

The objective questions yield higher coefficients (an average of .96 compared with .89) but the discussion questions are surprisingly objective, especially when we recall the relatively short length of the tests. The coefficients varied somewhat from reader to reader, particularly for

TABLE III

CORRELATION BETWEEN EACH READER AND THE AVERAGE FOR FOUR QUESTIONS OF EACH TYPE

	Reader								
	1	2	3	4	5	6	7	8	Average
Objective (4 questions) Reader . average	.98	.97	.96	.93	.97	.95	.97	.94	.96
Discussion (4 questions) Reader . average	.87	.92	.84	.87	.90	.90	.93	.85	.89

the discussion questions, where the range is from .84 to .93, but they are based upon a small number of cases and the variations are probably chance ones.

RELIABILITY

The measurement of reliability, the consistency of student performance on different samplings of the materials, will of necessity be attenuated by any subjectivity in the scoring process, so that the problem here is the practical one of determining the reliability of the sampling as expressed in terms of the scoring process used. One might determine the reliability when the sampling varies but the reader remains constant or when both sampling and reader vary. The more practical concern is with the reliability of different samplings for the same reader, and since high agreement among readers has been found, the study was confined to this aspect.

The Spearman-Brown method was used to determine reliability. In order that they might be comparable, the total, the discussion, and the objective parts were stepped up to the same length, 10 questions. Such a procedure equates the types as to length, but not necessarily as to testing time. Table IV presents for each reader's marks the stepped-up coefficients for the total, for the discussion, and for the objective questions, and in addition stepped-up coefficients for the combined eight readings for one-half with the other half. For these last coefficients the objectivity is at a maximum, and they may be interpreted as indications of the inherent reliability of samplings of the two types, although they have little practical significance.

TABLE IV

STEPPED-UP COEFFICIENTS OF RELIABILITY FOR TOTAL EXAMINATION, FOR OBJECTIVE QUESTIONS AND FOR DISCUSSION QUESTIONS

	Reader									
	1	2	3	4	5	6	7	8	Average	Average of All Readings
Total examination (stepped-up to 10 questions)	.82	.82	.80	.80	.90	.89	.81	.84	.84	.89
Objective questions (stepped-up to 10 questions)	.87	.86	.81	.89	.90	.90	.91	.91	.88	.90
Discussion questions (stepped-up to 10 questions)	.82	.81	.74	.70	.79	.87	.75	.71	.77	.85

The average coefficient of the several readers for the total examination was .84, for the discussion questions stepped-up to the same number was .77, and for the objective questions was .88. The discussion type seems to be less reliable than the objective, and this is still true although not so much so when we consider the reliability of the average of all readings. The reliability is satisfactorily high, and probably compares favorably with that of the usual teacher made new-type test.

CORRELATIONS WITH A NEW-TYPE TEST

Several studies found in recent literature report a very high correlation between the essay examination and new-type objective tests. Paterson (5) reports an average r of .52 for several correlations between essay and objective tests, which when corrected for attenuation (using the average reliability which he reports for making the correction) becomes .88; Corey (1) reports a coefficient corrected for attenuation of .93; Peters and Martz (6) report average uncorrected coefficients as follows: essay and true-false .74, essay and multiple-choice .78, essay and completion .76; Sims (10) reports a corrected coefficient of .99; and Eurich (2) reports corrected coefficients as follows: essay and completion 1.20, essay and multiple-choice .97, essay and true-false .89. However, Eurich also reports for another examination (in Educational Statistics) corrected coefficients between types mentioned above of .62, .67 and .66, respectively. This finding suggests decided differences among examinations. We have pointed out that a large number of elements in the typical essay examination are objective, that is, are of the same type as those found in new-type tests. To what extent is the high correlation between the two types due to the presence of these "identical elements"? This question is a fundamental one, for in spite of all improvement the essay examination can probably never give as good measurement as does the new-type test. If then the two types measure the same thing, one would recommend the new-type for efficient and accurate measurement. On the other hand, if it can be demonstrated that the essay examination, or certain types of essay questions, measure abilities different from those measured by the new-type tests, and abilities that we consider desirable outcomes of instruction, then the essay examination deserves a place in our testing program.

The 33 students who were used in this experiment were given a new-type test covering the same subject matter. This test consisted of 40 true-false statements and 32 completion items and had a reliability co-

efficient of .81. Although the two tests were designed to cover the same subject matter, no attempt was made to test on the same elements. In fact the objective test was prepared the year before.

Table V presents the coefficients of correlation, both raw and corrected for attenuation due to unreliability, between the new-type test and the average of the eight reader's scores on the total essay, the objective portion, and the discussion portion. The correlation between the objective questions and the discussion questions is also reported.

TABLE V

CORRELATIONS BETWEEN A NEW-TYPE TEST AND THE TOTAL EXAMINATION, THE OBJECTIVE QUESTIONS AND THE DISCUSSION QUESTIONS

	r raw	r corrected for attenuation
New-type with total	.83	.97
New-type with objective	.84	1.01
New-type with discussion	.49	.64
Objective with discussion	.40	.53

The indication is that the objective questions and the new-type test measure the same thing, while that which is measured by the discussion questions is positively, but not closely related to it. The correlation between the new-type test and objective questions is very high ($r = 1.01$), but the discussion questions correlate with neither very highly ($r = .64$ and .53, respectively). This point obviously needs further experimentation, for it may be the determining factor in the final disposition of the essay examination.

SUMMARY

An essay examination consisting of six objective (simple-recall and short-answer) questions and four discussion questions was given to a group of 33 students and the papers were scored by eight readers all using the same scoring rules. The rules involved using a scoring key for the objective items and rating or sorting the discussion questions. The scores assigned to the set of papers by each reader were converted to percentage grades by graphing the relation between scores and grades, assuming the passing score to correspond to the passing grade and the maximum possible score to correspond to the maximum possible grade.

When so scored and graded:

1. The variation among readers was slight. The average range for the grades given a paper was 4.4 points, the average probable error was 1.0 points.

2. The objectivity of the examination was high. The average correlation between a reader's marks and the average of all reader's marks was

.97. The average coefficient between the marks of paired readers was .94. The average coefficient between a reader's marks and the average of all reader's marks for four of the objective questions was .96. The average coefficient between a reader's marks and the average of all reader's marks for the four discussion questions was .89.

3. The reliability was very satisfactory. The average reliability coefficient for the total examination was .84, for the objective questions stepped-up to the same number was .88, and for the discussion questions stepped-up to the same number was .77.

4. The two types of questions (objective and discussion) seemed to measure different things. The correlation between them, when corrected for attenuation, was .53. The correlation between the objective questions and a new-type test over the same subject matter was 1.01 when corrected for attenuation, while the correlation between the discussion questions and the new-type test was .64 when corrected for attenuation.

REFERENCES

1. Corey, S. M. "The Correlation between New-Type and Essay Examination Scores," *School and Society*, XXXII (December 20, 1930), pp. 849-50.

2. Eurich, A. C. "Four Types of Examinations Compared and Evaluated," *Journal of Educational Psychology*, XXII (April, 1931), pp. 268-78.

3. Holzinger, K. J. *Statistical Methods for Students of Education.* Boston: Ginn & Company, 1928.

4. Kelly, F. J. *Teachers Marks.* Contributions to Education, No. 66. New York: Columbia University, 1914.

5. Paterson, D. G. "Do Old and New-Type Examinations Measure Different Mental Functions?" *School and Society*, XXIV (August 21, 1926), pp. 246-248.

6. Peters, C. C., and Martz, H. B. "A Study of the Validity of Various Types of Examinations," *School and Society*, XXXIII (March 7, 1931), pp. 336-38.

7. Ruch, G. M. *The Objective or New-Type Examination.* New York: Scott, Foresman and Company, 1929.

8. Shriner, W. O. "The Comparison Factor in the Evaluation of Examination Papers," *Teachers College Journal*, I (January, 1930), pp. 65-74.

9. Sims, V. M. "A Simple Technique for Converting Test Scores to Percentage Grades," *School and Society*, XXXIII (May 2, 1931), pp. 596-98.

10. Sims, V. M. "The Objectivity, Reliability and Validity of an Essay Examination Graded by Rating," *Journal of Educational Research*, XXIV (October, 1931), pp. 216-23.

11. Sims, V. M. "Essay Examinations Classified on the Basis of Objectivity," *School and Society*, XXXV (January 15, 1932), pp. 100-102.

12. Sims, V. M. "Reducing the Variability of Essay Examination Marks through Eliminating Variations in Standards of Grading," *Journal of Educational Research*.

FACTORS INFLUENCING THE SELECTION OF ASSOCIATES

MAY V. SEAGOE
Stanford University

EDUCATION in the broadest sense is growth guided toward socially desirable ends. This social development is of fundamental importance. Education can control social development only through a knowledge of its processes, and such knowledge must begin with the simplest forms of social relationship. The present study analyzes one of these forms, namely, child friendship.

I. PROBLEM AND METHOD

This study has been limited to factors which influence children from the third grade through the eighth in their selection of associates. The term "selection" includes chance as well as purposive factors. Two questions are investigated: (1) To what extent is the child an active agent in selection, and how much is he limited by propinquity? (2) Is similarity or dissimilarity the more important in the selection of associates?

A. *Preliminary Investigation.*—Before the study proper was attempted a preliminary study was made to define the problem and to point out fruitful approaches for more minute analysis. A questionnaire covering possible items of importance was submitted to 72 boys and 70 girls representative of grades five to eight, inclusive.

The first problem was to determine the most valid criterion of intimacy. The child was asked to name his choice for best friend, for work companion, for social companion, and for confidant. Teachers were asked to name outstanding cases of friendship, and the results from the two sources were combined. The relative worth of the four criteria was judged on the assumption that the one presenting the highest degree of agreement with the rest of the choices made was the best indication of intimacy. On the basis of the findings it seems that either the best friend or the social companion is a better index than the work companion or the confidant. The social companion is used in the present study because of the specific nature of the situation pictured.

The second phase of the preliminary investigation consisted of an analysis of external factors leading toward friendship, *i.e.*, propinquity. On the basis of children's judgments of distance between homes the relationship was practically the same as that derived from the study proper. Data concerning length of acquaintance showed that in 89.4

percent of the cases the friendship had lasted more than one year; that length of acquaintance increases with age; that length of acquaintance increases somewhat with greater intimacy; and that length of acquaintance is to a large extent dependent on continued geographic propinquity. Length of acquaintance is neglected in the study proper because of the difficulty of objective analysis and because of its dependence upon propinquity.

Opportunity for association at school as judged by the grade attained and the room in which the child was enrolled showed a relationship high enough to justify its inclusion in the study. A rough summary of chronological ages and of heights indicated that evidences of physical maturity were promising factors. A summary of intelligence quotients showed a similar relationship. A summary by sex showed a strong tendency for children to choose as friends members of the same sex, a tendency which increased somewhat throughout the age range studied. Eye-color was not included in the major investigation because its judgment is subjective and because the preliminary work indicated no definite trend either toward similarity or dissimilarity. The personal traits given greatest weight were cleanliness, sportsmanship, optimism, kindness, dependability, agreeableness, helpfulness, similarity to the friend, refraining from self-glorification, and external appearance.

B. *The Study Proper.*—The data for the study proper were gathered from 115 pairs of friends chosen from 823 children in grades three through eight of the Garvey district in Los Angeles county. The pairs consist of all children who chose each other in response to the test question: "Suppose you were going to a party and could invite just one person to go with you. Suppose it could be either a boy or a girl, but not a member of your family. Whom would you ask to go with you?" When the same question was repeated one month later, only 29 of the original 115 pairs again chose each other. On the assumption that the 29 represented a closer degree of intimacy, they were treated separately as Group I. The remaining 86 constituted Group II.

It is significant that in all of the 115 pairs there were no choices between members of opposite sex. The cases consisted of 55.2 percent girls and 44.8 percent boys. It might have been preferable to summarize responses for boys and girls separately, but since sex differences are supposedly slight for the major part of the age range included and since none of the choices were between members of the opposite sex, responses have not been differentiated.

The study was planned to cover the factors listed in Table I. The child was asked for items of common information and for judgments concerning his fellows; all data for which more reliable sources were available were compiled from such sources. Propinquity was approached through home location and through school grade; physical similarity was attacked through sex, chronological age, height, weight, and weight-height ratio; mental similarity was studied through intelligence quotient and mental age; and similarity in personal characteristics was judged through pupil opinions.

TABLE I

PLAN OF THE STUDY

Factor Studied	Method of Attack: Item Needed	Method of Attack: Source	Statistical Treatment
I. Propinquity:			
A. Home	Address	Child	Get average distance between homes.
	Map	Principal	Express home location numerically. Compute r.
B. School	Grade	Child	Correlate grade positions
	Teacher	Child	Compute percents in same grade and room.
II. Physical similarity:			
A. Sex	Statement	Child	Compute percents.
B. Chronological age	Register	Teacher	Compute r.
C. Height	Health record	Nurse	Compute r.
D. Weight	Health record	Nurse	Compute r.
E. Bodily type			Transmute height and weight to metric units. Compute ratios. Correlate ratios.
III. Mental similarity:			
A. I. Q.	Test record	Counselor	Compute r.
B. Mental age			Derive from C.A. and I.Q. Compute r.
IV. Similarity in personal characteristics:	"Who Is It?" Test	Child	Convert into scores. Average for each child in each trait. Compute r for each trait and for composite.

Four forms were used. The first form was a pupil choice report on which the pupil recorded his name, address, grade, teacher, and sex, then gave the same items as well as he could for the person he chose. Information for the person chosen was verified from school records. The second form was the "Who Is It?" Test. This device asks the child to name a person in his room who is like a given description. The number of times a child is named by his classmates constitutes his score in each characteristic. Twelve traits emphasized in the preliminary investigation were used as a basis for the test. Either a positive or a negative score was possible in each trait. The third form was submitted to the child one month after the first and again asked his choice of social companion. The fourth form was a summary blank bringing together the data from the other forms and supplementing it with height, weight, age, and in-

telligence test data. The intelligence quotients were in most cases averages of two group tests.

With the records complete, statistical work was begun. The homes were plotted on a map, air-line distances between pairs were measured, and a "home index" was devised. In order that the relationship in home location might be expressed in the same terms as relationships in other factors it was necessary to find a means of expressing each location numerically in order that the figures might be correlated. The purpose was not to glorify the correlation technique, but merely to reduce data to comparable terms. The supposition was that if pupils live near each other they will live approximately the same distance from any reasonably distant point. A reference point just outside one corner of the map was therefore selected and a right angle was described from the point along two sides of the district. The home index was the product of the shortest distance from the home in question to each of the reference lines, *i.e.*, the area of the rectangle thus described. A product rather than a linear measure was used to minimize the factor of direction.

The index is not, of course, infallible. The same product might be derived from more than one location. The use of the index was validated in two ways. (1) A second reference point at an adjoining corner was selected to see whether the correlation depended upon the particular point chosen. The coefficient from the first point was .874 ± .031, and that from the second was .744 ± .058. Both are statistically significant and within two probable errors of each other. (2) Using alphabetic pairing a random grouping was obtained. The coefficient for random pairs was .056 ± .067. On the basis of the present data, therefore, the home index appears to be statistically sound.

Weight-height ratios were derived using the formula

$$\frac{W}{H^3} \times 100$$

in metric units. Mental ages as of the date of the study were obtained by multiplying the intelligence quotient by the chronogical age. The results of the "Who Is It?" Test were tabulated; the number of times the child's name was mentioned by his classmates as possessing or lacking each trait was found, and the algebraic sum recorded as his score. The sum of all characteristic scores formed the composite. Means, sigmas, and coefficients of correlation were computed for both groups for the home index, grade, chronological age, height, weight, weight-height ratio, intelligence quotient, mental age, the twelve personal traits, and for their composite.

The average distance between homes for all pupils studied was obtained by dividing the district into sixteen blocks, distributing the cases by blocks, measuring the distance from each block to all others, weighting each distance by the number of cases, and dividing the total number of weighted distances by the total number of homes.

II. RESULTS

The means and standard deviations were valuable only in showing that the subjects used were representative of the general population of the grades included. All measures were near the norm for the age in question.

A summary of correlations is presented in Table II. The coefficients for Group II are probably more reliable on the whole because the number of cases involved was considerably larger.

TABLE II

COEFFICIENTS OF CORRELATION

Factor	Coefficient of Correlation	
	Group I	Group II
Propinquity:		
Home index	.874 ± .031	.766 ± .032
Grade	.959 ± .010	.949 ± .007
Maturity:		
Chronological age	.866 ± .031	.803 ± .025
Height	.422 ± .126	.543 ± .055
Weight	.459 ± .108	.634 ± .047
Weight-height ratio	— .149 ± .135	.010 ± .116
Intelligence:		
Mental age	.675 ± .078	.779 ± .031
Intelligence quotient	.513 ± .105	.386 ± .067
Personal characteristics:		
Beauty	— .097 ± .123	.056 ± .073
Athletic ability	.343 ± .110	.346 ± .064
Courtesy	.167 ± .121	.299 ± .066
Ability to interest listeners	— .095 ± .123	.116 ± .071
Cleanliness	.349 ± .110	.468 ± .057
Sportsmanship	— .116 ± .124	.234 ± .068
Amiability	— .027 ± .124	.153 ± .071
Kindness	— .243 ± .118	.112 ± .071
Dependability	— .242 ± .118	.145 ± .071
Lack of quarrelsomeness	.210 ± .120	.121 ± .071
Helpfulness	.374 ± .108	.017 ± .073
Lack of snobbishness	.277 ± .116	— .112 ± .071
Composite	.140 ± .123	.320 ± .065

The measures of propinquity show the greatest relationship. The mean distance between homes was .26 miles for Group I and .33 miles for Group II. The average distance between unselected homes was .92 miles, or about three times that between the homes of associates. Sixty-two and nine-tenths percent of Group I lived less than .20 miles apart. Greater geographic proximity appears to be the only respect in which Group I is consistently more closely related than Group II; it seems to be a potent factor in the continuation as well as in the formation of friendships. Pro-

pinquity as measured by the home index also showed a close relationship.

Propinquity at school was of considerable importance. About 75 percent of both groups were in the same grade and an equal percent in the same room. The correlation between the grade attained by associates was high.

The correlations between measures of physical maturity were also significant. When age was held constant in the weight-height ratio, however, the correlation disappeared. It is physical maturity rather than bodily type which produces the correlation.

The correlation between mental ages was relatively high. When age was held constant in the intelligence quotient the coefficient was lower, but still significant. There appears to be a reliable relationship in mental capacity as well as in mental maturity.

Finally, the twelve personal characteristics showed some relationship. Data for Group I can be interpreted only in the light of agreement with Group II due to the small number of cases and to the relatively low coefficients. Correlations more than four times their probable errors in Group II were found for athletic ability, courtesy, cleanliness, and total score.

III. INTERPRETATION

The results seem to show two tendencies. (1) Propinquity as a force limiting friendship is given emphasis. (2) All the data on similarities shows that likeness is of greater importance than difference in social attraction.

The factor of propinquity may function in two ways. First, propinquity limits opportunity for acquaintance. Such limitation is conditioned (a) by a chance element which is shown in measures of propinquity at home and at school, and (b) by a social element which is evidenced in measures of intelligence and of personal characteristics. The existence of the social element is based on the assumption that there is a differentiation of social background among occupational and other types of social groups, and that the child's personal characteristics are to some extent moulded by his social background. Similar homes produce similar child attitudes, which in turn result in attraction between the children. The child, being influenced in his selection by similarity, selects for the most part companions whose social heritage is similar to his own. Courtesy and cleanliness are two characteristics in which this tendency is shown in the present study.

Second, propinquity limits the opportunity for continuation of contact, and it is continued contact which keeps friendship alive. Contact leads to habituation. Expressing this deterministic view, less technically, we do not deliberately institute a search for persons of a set pattern for our friends; we are limited by the breadth and the depth of our contacts. The educational implication is obvious; throw the child with the sort of persons with whom you wish him to be intimate; they will help form his tastes, and with the tastes an attachment to persons having those tastes.

Within the limits set by propinquity, however, we select associates according to the satisfyingness of our contacts, and according to the pleasure we obtain through our associates. Propinquity marks the boundaries within which we may choose; similarity conditions our choices within those boundaries. We enjoy the companionship of persons of the same age and physical maturity, of the same intelligence, and of the same disposition. The desire for response seems to be fundamental in human nature; perhaps our selection of similar associates is related to the desire for response. Similar persons are most capable of responding to all our moods; an inferior would not be sufficiently responsive, and a superior would too often offend by surpassing us and thus wounding our self-esteem. Similarities were most obvious in the case of physical maturity in the present study, but it is possible that if intelligence and other personal characteristics could be measured with the precision used in physical measurements the relationship there might be equally high.

The study is, however, only a suggestive attack upon two factors in friendship. The age-range should be increased. The factor of age should be equated in all measures. Better measures of relationship between homes and better measures of personal characteristics must be devised. A different method of determining the most intimate friend should be used to check the results presented. Suggestions for further study in this field include: (1) An analysis of differences between sexes, between age-groups, between mental levels, and between varying degrees of propinquity in the extent to which each of the factors suggested above operates. (2) An evaluation of the personal characteristics to determine intercorrelation, correlation with intelligence, and correlation with age in order to ascertain the degree to which personal characteristics are dependent upon other factors. (3) Further studies of propinquity showing the extent to which propinquity colors relationships in age, grade, etc., the extent to which density of population increases the more consciously selective factors in friendship, the numerical extent of the social contacts of

the average child, and the extent to which Sunday-school, work, social contact of families, and other similar conditions influence the choice of companions. (4) A study of the constancy of child friendships and the factors involved. (5) A study of leadership using a similar technique to determine the extent to which friendship and leadership are related. (6) An experiment controlling the factor of propinquity by rotation and analyzing the consciously selective factors which appear.

IV. CONCLUSIONS

In summary, the investigation shows that:

1. Measures of propinquity show the highest degree of relationship.
   a. Correlations for home location are in the seventies and eighties.
   b. Mean distance between the homes of associates is .26 miles compared with .92 miles for unselected pupils.
   c. Correlations for grade approximate .95.
   d. About 75 percent of the pairs of associates are in the same grade and room.
2. Measures of physical maturity correlate significantly.
   a. Age, height, and weight give coefficients ranging from .86 to .42.
   b. When age is held constant, the relationship between weights and between heights disappears, showing the correlation to be a function of maturity.
3. Measures of mental level and of mental capacity correlate significantly.
   a. Mental ages correlate .67 and .78.
   b. When age is held constant in the intelligence quotient the correlation is lower but still significant, *i.e.*, .38 and .51.
4. There is a slight but significant relationship between associates in the personal characteristics of athletic ability, cleanliness, courtesy, and total score in socially desirable traits.

Evidence is therefore presented in support of two principles, namely: (1) Propinquity is a strong factor in the selection of associates among children. (2) Similarity is more important than dissimilarity in the selection of associates.

SELECTED BIBLIOGRAPHY

1. Almack, John C. "The Influence of Intelligence on the Selection of Associates," *School and Society*, XVI (November 4, 1922), pp. 529-530.

2. Bonser, F. G. "Chums: A Study in Youthful Friendships," *Pedagogical Seminary*, IX (June, 1902), pp. 221-236.

3. Challman, Robert C. "Preschool Friendships," *Child Study*, VIII (December, 1930), pp. 106-107.

4. Cheek, William Virgil. *Social Aspects of Intelligence*, Stanford University Thesis, 1924.

5. Furfey, Paul H. "Some Factors Influencing the Selection of Boys' Chums," *Journal of Applied Psychology*, XI (February, 1927), pp. 47-51.

6. Warner, M. LaVinia. "Influence of Mental Level in the Formation of Boys' Gangs," *Journal of Applied Psychology*, VII (September, 1923), pp. 224-236.

7. Wellmann, Beth. "The School Child's Choice of Companions," *Journal of Educational Research*, XVI (September, 1926), pp. 126-132.

8. Williams, Paul E. "A Study of Adolescent Friendships," *Pedagogical Seminary*, XXX (December, 1923), pp. 342-346.

STATISTICAL DETERMINATION OF THE SOCIAL SCIENCE CURRICULUM

EDWARD O. SISSON
Reed College

THE advance of the social sciences in the curriculum during the last twenty years has been nothing less than phenomenal, particularly in the secondary school where this subject has won the paramount place so long held by the so-called classics. The subject-matter in this field is plainly of prime importance in training for life and citizenship; it is also so rich and varied as to make the problem of curriculum-building peculiarly formidable. It was inevitable that the new technique of measurement and statistical procedure should be called into service in the selection and organization of material; this technique has proved its value in other fields, notably spelling and arithmetic, and will doubtless give aid in the social sciences. This article is written with the view that such statistical research is desirable and indeed necessary, but is subject to peculiar hazards, some of which it is the purpose of this paper to discuss.

We may begin with one of the more ambitious projects in the field, namely, Washburne's study published in 1923 in the Yearbook of the National Society for the Study of Education, Part II. This study presents a ranked list of 1500 "persons, places, dates, and events" (although strictly speaking, there are no dates as such in the list). The list is to "form an objective basis for a fact course in history and geography". The 1500 topics are the boiled-down quintessence of some 80,000 "mentions" gathered from periodicals and books over a period of six years; each item in the final list is scored on the basis of its mentions, and the items under each score are ranked among themselves by a refinement of the statistical process.

The scores run from 103 down to six, items with scores below six having been discarded. The report then declares that "we thereby obtained a strictly objective ranking of all our items . . . It shows, therefore, the relative importance of these items from the standpoint of enabling children to read intelligently." A list of items taken from the extremes of Washburne's list is given below to illustrate the problem under discussion.

Some Items with Score 6	*Some Items with High Scores*	
Pan-American (Federation, Congress, etc.)	Greece	78
Continental Congress (American)	Switzerland	69
Interior Department of U. S.	Rudyard Kipling	56
John Marshall	Persia	50
Civil Service	W. J. Bryan	47
John Dewey	Los Angeles	36

This list could be continued almost *ad libitum* without diminishing the violence of the contrasts. Astonishing juxtapositions also occur: under score 7 appear these two sets of items:

Prohibition	Cornwall
Herodotus	Kaffir
Valley Forge	Howard Pyle
Pasteur	Bessarabia
Alexander Graham Bell	Strand (London)
Versailles Peace Treaty	Grand Central Station (N. Y.)

Scoring 9, that is fourth from the bottom, are American Expeditionary Force, Peace Conference, Central Powers (World War), Magna Charta, Cabinet (U. S. President's); under the same score there is also a mass of trivialities. Scoring 18 we find the American Federation of Labor, and the Constitution of the United States; but also Tibet, San Diego, Solomon (King), Central Park, and Bronx. The Supreme Court of the U. S. climbs to a score of 29, along with Rochester, N. Y., and Henry C. Lodge.

The scoring of persons is no less surprising: Rudyard Kipling, 56; Christ (including Jesus), 50; and J. Daniels, 24. Henry Van Dyke outranks Scott and Goethe. Martin Luther, Daniel Webster, Cicero, Constantine, and Lenin score 15; but also do D'Annunzio, Thomas Hardy, Champ Clark and Huerta. Under 8, come Henry George, Henry Clay, John Quincy Adams, Wm. M. Lorimer, William J. Locke, and Joseph Pennell. Obscured in the lowest rank are John Brown, John Stuart Mill, Pericles, Peter the Great, and Patrick Henry; with them are placed Lee Schubert, Roger M. Sullivan, and George M. Cohan.[1]

[1] Another study based upon counting mentions offers a list of 59 persons: the list includes Buchanan, Burr, Bryan, Harrison, McKinley, Shaw, Taft, Tyler; but fails to include Columbus, Charlemagne, Burke, Emerson, Frederick the Great, Pasteur, Socrates, Jesus, James Watt, or Walt Whitman. (Earle Rugg. *Curriculum Studies in the Social Sciences and Citizenship.* p. 53.)

A list of 18 cities takes in Glasgow, Manchester, Stockholm, but omits Rome and Geneva. (*Op. cit.* p. 38.)

Two methodological errors, both of vital importance in relation to the problem of the curriculum, seem to stand out in the study: first, words are counted rather than ideas or concepts. Item 1, with its score of 101, is "America" (meaning U. S.); it is followed lower down by "United States," "South" (U. S.), "West" (U. S.), "East" (U. S.), "North" (U. S.), and "Middle West" (U. S.). Item 2, "England," is followed by "Great Britain," "Ireland," "Scotch, Scotland," and far down the list, "British Isles," "Britain" (ancient and modern), and "Wales." Curiously enough *British Empire* is not to be found, although high up are "Canada," "India," "Egypt," and "Australia."

The second main error lies in the counting of mentions without weighing them. Certainly the list does show one thing, namely, the relative number of times the various items were mentioned in the material covered by the study; but it would seem to be a great leap from this to the conclusion that frequency of mention represents importance in a social science curriculum. It would not appear statistically defensible to include in a series of sums a series of unequal units such as are merely counted in this investigation. The equality of units is the mathematical prerequisite for any counting; certainly, absolutely equal units do not exist in the world of reality, at least within our control; and our best measurements are approximations. Yet, unless the units may properly be taken as equal, the mathematical procedure is invalid, and the data arrived at are meaningless. It may be urged that the quantification of mentions is inherently impossible; if this is true, it would seem then that the counting of mentions must be given up as a mode of determining importance. It seems clear, however, that it is not true, for certain studies do undertake to quantify mentions, one method being to draw material only from the writings of experts, and the results seem to afford support for the procedure.[2]

Even with the aid of methods which tend to select only relevant and important mentions the process of ranking seems hazardous. The Billings study first uses only expert matter, the writings of "frontier thinkers" in the field; it also adopts an elaborate scheme of determining, first a set of generalizations, next an organized group of central themes, and finally extracting from these a ranked list of concepts, 505 in number. Again we have rankings which seem quite unjustified. Consider the following

[2] Hockett, J. A. *A Determination of the Major Social Problems of American Life* (1927); Billings, Neal. *A Determination of Generalizations Basic to the Social Sciences* (1929). Five studies reported in E. Rugg, *op. cit.*, Table XVIII, p. 162, exhibit other methods of selecting only relevant and important mentions.

exhibit of high and low scores (the top score is 349, the lowest 2, with 101 items; the median 6):

Homogeneity	91	Constitution (political)	3
Specialization	70	Cost of living	
Personality	44	Economic wellbeing	
Ability	39	Freedom of speech	
Sympathy	32	Markets	
Energy	30	Statesmanship	2
Impulse	12	Finance	

This list could be extended greatly without materially diminishing the discrepancy between the high and low items. It is true that the contrast here is not between great and small items, for the items in the high list are all matters of profound moment, but they are hardly of a nature to afford guidance in the formation of a curriculum; the low-score list, on the contrary, is composed entirely of social science topics of the very first rank, yet the mean rank of these topics would demote them to meager consideration or even hazard their omission.

It would seem that the trouble here is at least in part *verbalism:* the first column above includes words which tend to figure abundantly in any discussion of human affairs, but hardly one of them really denotes a social science concept; while the second column does present a group of such concepts. The top item in the list, with its supereminent score of 349, is "thought (or men's thoughts, motives, desires, etc.);" the sixth item is "life (in general) or life activities;" item 15 is "economic conditions in general (economic life)."[3] Not a few items among the high scores are vague or general in the highest degree: such as "supply," "association with others," "institution," "happiness," "human history," "knowledge," "philosophy," "men's spirit (attitude)," "ambition," "intelligence," "emotion (men's emotions)"; all these and many like them are in the top quarter of the list with scores of 23 or over. It is hard to escape the view that all these items are indeed vital and significant aspects of life, but very far from being either social science concepts or suitable topics for the curriculum.

For a third case of the use of ranking we turn to Table CVIII in Rugg's study referred to above. The table is entitled: "Comparison of Twenty Topics Judged Most Important on Criteria of Life Values and Actual Rank (proportionate emphasis) in Representative Textbooks."

[3] Curiously enough this vast theme is outranked by "price," "geographical environment," "location of political power," and "migration," all important topics yet not conceivably of more importance than a topic which seems to comprise the whole science of economics.

The table sets out the top twenty topics in six lists; the first five lists are from different sources and by different hands; the first is based upon platforms of the two major political parties, and the other four are based upon periodical literature with some inclusion of books. Each of these five studies is more fully reported earlier in the book. The sixth column is headed "Textbooks" and is apparently intended to be broadly representative, but the writer has been unable to trace its definite derivation. The table makes no claim to authoritative determination of relative importance of topics; however, it serves to exhibit the hazard of ranking.

Particular items range from high to low; each list contains items not found in any other and each omits items prominent in others. "Education" runs from fourth to twentieth place and is absent in two lists; "world-order" and other terms closely synonymous rank at or near the top in three lists and are absent in the other three. "Bolshevism" stands fourth and eighteenth in two lists and is absent from the rest; it should be added that "radicalism" is listed tenth in a third list, and "socialism" is eleventh in the list which places "bolshevism" fourth. "Immigration" ranks third, fifth, tenth and eighteenth, with two absences. "Prohibition" gets into two lists only, ranking twelfth and eighteenth.

Verbal identity is rare between the lists; four terms occur four times each: education, recreation, immigration, and labor. Three occur three times: finance, taxation, and commerce; ten items occur twice. Thus we have a total of 92 different terms in the six lists with their 120 items. This, of course, throws the problem back to concepts; but here, as elsewhere in ranked lists, this undertaking is too formidable for the framers of the lists; it is easy to count words, but also futile.

The six lists conflict not only in details but also in the largest phases. If we order the topics under certain large general heads, such as politics, industry and labor, commerce, finance, social problems, foreign relations, we find the greatest discrepancy in rank and stress: politics, commerce, finance, and social problems are absent each from one list; finance is absent from two. The author himself presents a final list of 24 topics, but with no indication of the method of selection. Here, again, strange things occur; social problems with a total of 13 mentions in the tabulated lists does not get into this final list, yet room is found in it for *canals* and *pensions*.

Many other examples of ranked lists are available in the literature of education. While later methods of collecting data tend to correct the gross inequality of mentions referred to above, no means has yet been effectively used to escape the mere verbalism of those lists.

A TECHNIQUE FOR MEASURING THE ABILITY TO EVALUATE OBJECTIVE TEST ITEMS[1]

J. R. GERBERICH
University of Arkansas

THIS report presents an experimental development of a test technique for measuring the ability to evaluate objective test items. Courses dealing with the new-type classroom test should be designed not only to equip the student with a knowledge of the principles underlying the construction of objective test items but also to develop a critical attitude in the evaluation of such items. It is probable that more emphasis is usually placed upon factual knowledge in measuring the outcomes of such courses than upon the more intangible outcomes which are dependent upon but which do not necessarily accompany factual knowledge.

A special mimeographed examination, consisting of fifty true-false and eighteen multiple-response items, was given to the writer's 1931 summer session class in "Construction and Validation of Objective Tests" at the end of the unit dealing with item construction and evaluation. The class consisted of forty junior, senior, and graduate students, with the last group making up approximately half of the total enrollment. No time limit was set for the test except that imposed by the fifty-five-minute class period.

The usual types of true-false and multiple-response items were utilized differently in this test than in the usual objective examination. The student was asked to evaluate objective items from the standpoint of a teacher contemplating their use in a classroom test. Subject matter from geography, literature, educational measurement, and several other sources was employed in the items to be evaluated. The subject matter chosen was of such an elementary nature that each student evaluating the test elements was almost certain to possess adequate factual knowledge for their interpretation. Furthermore, as the directions for the test pointed out, the truth or falsity of the items (from the standpoint of the pupil who might encounter them in the usual test situation) mattered little, except as a possible supplementary basis for the student's evaluation of a few items.

The test elements were of two types: (1) Good and satisfactory items, not subject to major criticism as to form or style, (2) Ambiguous items, introducing factors which would tend to confuse the pupil unnecessarily

[1] Research Paper No. 291, Journal Series, University of Arkansas.

in the usual testing situation and which would, therefore, measure his knowledge of the subject matter plus his ability to puzzle out the intent of the teacher. The student was told to indicate his reactions by encircling the "Good" or "Poor" following each item, according to whether he felt it to be satisfactory or ambiguous. He was instructed to evaluate the method of testing, form, and style of the items, and to disregard the worthwhileness or importance of the facts or principles upon which they were based.

The second aspect of the student's response was presented in the directions by the following paragraph:

"After you have indicated your reaction to an item, proceed as follows:

(1) *If you marked it "Good," go on immediately to the next item, without writing anything on the line indicated for comment.* (2) *If you marked it "Poor," write on the line indicated for comment your reason for thinking the item should be eliminated or revised.* Use just enough words to make your meaning clear; do not use complete sentences. For example, "Double-negative," "Incorrect grammar," "More than one correct response," and the like are satisfactory, but "Ambiguous" would be unsatisfactory because of being indefinite.

A comparison of four different scoring methods was made, in order to determine the one having the highest reliability and validity. In each scoring method it was assumed that a student who correctly marked an item as "Poor" but who gave no logical reason for such a choice had made almost a pure guess or had possessed only a vague idea that the item was not entirely satisfactory. Therefore, such responses were given no credit in any of the scoring methods.

Four scores were determined for each student. The first two were obtained from unweighted and weighted correct responses, while the last two were derived in the same manner as the first two except that correction was made for chance. The second method, which allotted one and two points respectively to each "Good" and "Poor" item correctly answered and which made no correction for chance, appeared to be the most satisfactory. Objectivity of scoring was indicated by a correlation coefficient of .985 between scores assigned independently by the writer and another scorer. A reliability coefficient of .930 indicated the test to be highly reliable.

The validity of the items making up the special examination was determined by an item analysis. The papers were classified into three

groups on the basis of the term grade, A and B students in the first, C students in the second, and D, E, and F students in the third. The percentages of students correctly answering each item were computed for the three groups. A few representative items are given below, together with the percentages of correct responses given by students of the high and low groups.

Item Number	True-False Items	Correct Response	Percentages of Correct Response	
29	It appears evident that the new-type examination has higher reliability than the essay examination. (Use of "appears evident" conditions response)	Poor	High Low	92 0
23	Tennyson wrote "Idols of the King". ("Idylls", not "Idols"—intentionally confusing)	Poor	High Low	77 27
38	Milton wrote "Il Penseroso".	Good	High Low	100 64
2	Sufficient data is available to show the unreliability of the essay examination. ("Data is"—incorrect grammar)	Poor	High Low	92 27
49	Robert Burns was known to be a misogynist. (Difficulty of key word—"misogynist")	Poor	High Low	85 45
17	In the development of testing techniques, less attention has been given to the measurement problems of the high school. (Comparison incomplete)	Poor	High Low	38 0
32	"The Prisoner of Chillon" was written by Hawthorne.	Good	High Low	85 55

Multiple-Response Items

15	The 1931 Nobel Prize in Literature was awarded to (1) William Shakespeare, (2) Sinclair Lewis, (3) Oliver Goldsmith, (4) H. L. Mencken, (5) Carl Sandburg.	Poor	High	92
			Low	36
	(Some obviously incorrect — not contemporaries)			
13	Memphis is an important city in: (1) Oklahoma, (2) Louisiana, (3) North Carolina, (4) Kansas, (5) Tennessee.	Good	High	85
			Low	64
11	The Orinoco is a: (1) Lake, (2) Ocean, (3) River, (4) Mountain, (5) Island.	Poor	High	69
			Low	36
	(Use of "a" eliminates some responses)			
8	Springfield is a city in: (1) Missouri, (2) Oklahoma, (3) Massachusetts, (4) Florida, (5) Illinois.	Poor	High	54
			Low	36
	(Several correct responses)			

The discriminative power of the examination as a whole is shown by the forty-four out of fifty true-false and thirteen out of eighteen multiple-response items which showed a greater or less degree of discrimination between students of the three ability groups. Forty-four of the sixty-eight items were intended to present some form of weakness, while the remaining twenty-four were essentially satisfactory. The items showing the greatest discriminative power were mainly those presenting some form of weakness, so that the major testing burden was borne by the "Poor" items. However, two-thirds of the "Good" items contributed to the discriminative power of the examination.

SUMMARY

A special 68-item examination devised to measure the ability to evaluate two types of objective test items was given to a University of Arkansas class which had just completed a course unit in construction and evaluation of objective test items. The true-false and multiple-response items were to be rated as "Good" or "Poor" on the basis of their form or style for use in objective classroom tests and without regard for the worthwhileness or significance of the facts involved. The "Poor" items fur-

ther required a brief statement of the reasons why they should be revised or discarded. The subject matter of the items was so selected that the students were almost certain to have a sufficient knowledge of the facts involved to serve as a background for interpreting the items.

The examination was experimentally scored by four methods. The optimum method, as selected on the basis of its reliability and correlation with a final examination and term grades, involved a weighting of one point for each "Good" and two points for each "Poor" item correctly answered. This scoring method had a reliability coefficient of .930.

EDITORIAL

RESEARCH IN AN EDUCATIONAL CRISIS

One of the most apparent aspects of the present educational crisis is the attack upon "fads" and "frills." While in general these "fads" and "frills" are those subjects most recently introduced into the curriculum—music, art, physical education, etc.—they may be any subject or group of subjects toward which, for various reasons, enmity may develop in the local community. The older subjects have the advantage of human inertia and of support by the tendency of each individual, through a process of rationalization, to believe that his own training was best, and therefore such subjects appear to be most secure. There is a tendency in some quarters to argue the case for the different school subjects. It is the writer's opinion that the lay public is not interested in nor greatly impressed by arguments for the different school subjects. If there is anything that is more apparent than anything else in the present educational crisis, it is the futility of mere argumentation. On the other hand, where carefully organized, unbiased facts have been presented they seem to have been generously respected. It has been many times asserted that we are living in a rapidly changing world; many of the newer subjects do have a place in this new world. A thing that is needed above everything else in the present educational crisis is a new evaluation of old and new curricula in light of changed conditions to determine not how much might be expended for education in an ideal society but the *essentials* of a good education in the new era.

A second point at which research can contribute to the dissolution of the present educational crisis is in the field of method. The fact that a method is "older" is, of course, no proof that it is better or more economical. This statement is, of course, equally true of "newer" methods. Among other tendencies, there appears to be in some communities an unthinking return to "older" and presumably thoroughly tried methods of instruction. There is, however, reliable evidence that some of the "newer" methods constitute real economies, comparable to those of our great industrial inventions. This should not be lost sight of in the press for economy. This situation makes method a worthwhile field for research and reasoned economy.

A third point at which research can contribute to the dissolution of the present educational crisis is in the field of "panaceas." Shortly we

shall probably witness such an outpouring of panaceas, each supported by plausible and to some people convincing arguments, as has not been witnessed for half a century. The schools and the public need to be protected from these "panaceas." There is probably no greater service, within certain limits, that research might render to the school and society than the evaluation of panaceas.

In brief, considering the tempo of the times, there has probably never been a time in American education when more careful thinking and research has been needed than at present: in the classroom and in the laboratory; and on the part of teachers, principals, supervisors, superintendents, and specialized research workers alike.

A. S. BARR

EDUCATIONAL RESEARCH IN THE CONTRASTING FIELDS OF CHARACTER AND TALENT

SOME good research work has been done in the field of socio-economic conditions in home life as related to the needed adaptation of school conditions to the differing opportunities of the children.

A related field for research but one that is more subjective in its connections, concerns the educational difficulties resulting from the lack of character forming activities for children in the modern homes and communities. These activities were once provided by the required services of all members in the chores and the responsible unions for work to do the many things now taken care of by modern mechanical appliances.

In a number of smaller school systems and more rural districts some of the school work is now made an integral part of directed work in the home or on the farm and in the community. This directed work has the form of laboratory experiences and includes personal responsibility for the processes as well as for the results. At the present time this relation of school work is generally classed as vocational training for an industry of mature life. Both the teachers and the pupils, however, find that the result of well organized and completed activities is an admirable training in character traits and habits.

Though vocationally limited these coördinations give definite illustration of the effect of completed work to illuminate, expand, and enrich the meanings of the particular experiences. Research is needed to determine how widely such coördinations between intellectual development and the attitudes, appreciated ends, and standards of conduct for self directed citizenship can be worked out for school and community activities not specifically vocational, except as social aims and standards are a vital part of all organized civic life.

Some very extensive experimentation has been initiated in a field assumed, possibly, to be closely related. The work of the Lincoln School of Teachers College, Columbia University, under Otis W. Caldwell, developed in the primary and intermediate grades a coördination between the nurture, through appropriate activities and constructive materials,

of individual talents and a broader nurture deepening an understanding of social values for a richer community life. The school provided materials and illustrative experiences to enable pupils to follow some of the steps by which primitive man advanced his understanding of how to get artistic results and gained the skill needed to make the instruments. Each pupil developed in some line of artistic expression and used the ability in some home or community situation. In the same school Harold Rugg experimented with material and method for social science studies. One of the outcomes, significant at least to Dr. Rugg, concerned the effect of the appreciation of beauty as a contributing value to social changes.

Experiments of this kind as well as the directly vocational organizations of school and trade experiences are not free from an implied and perhaps unconscious assumption that human thought and action run naturally along single track lines. Dr. Truman Kelley, in his book, "Crossroads in the Mind of Man," following Spearman's principle of the existence of independent mental factors or unit abilities, extends the possibility of breaking up a general factor of intelligence, and he provides a mathematical formula for the handling of correlations of data from five independent variables. By this he provides for mapping, as it were, the interconnections of various single track activities in the mind and thus shows how each, when crossing one or more of the others, may add or subtract some element of mental force for the final results.

This tetrad or the more powerful pentad difference formula is true in the physical field of force, when the action of force takes place in the mechanical use of machines. It is not shown to be true for the action of central energies such as gravity or chemism, where the central energy itself maintains motion or pressure and affects the direction of inertial movement by a constant factor which is independent of the fact that forces from impact or other external sources may cross the path and cause a change of the velocity and direction of motion.

This situation of theory is reported here to emphasize the need that both research workers and school supervisors shall seek the facts with regard to mental action as it embodies the energy for that fellowship which established great civilizations and fine realizations of justice, of loyal devotion to friends and to civic virtues, and of the harmonies of form, color, tone, and dramatic relations even in ancient times.

To the observer it seems possible that Spearman and Kelley are working along lines of investigation which parallel in a way the work of Galileo in determining the law of inertia for physical bodies of matter. He discovered that the inertia of mass moves it in a straight line without change in the velocity of motion (if not affected by gravity or by resisting forces) and therefore any force acting on mass is measured in terms of the product of the mass by the acceleration. This law of inertia made it possible for Sir Isaac Newton to conceive of the second law of matter as

a center of constant energy acting between masses. This action is defined as a force in direct proportion to the masses and in indirect proportion to the square of the distances apart. Gravity is the energy which holds masses, moving through space under the law of inertia, to a constant deflection from straight lines into orbits determined by the distances and by the other mass relations of all matter.

If this parallel is carried out for the energy relations between minds, then the immediate need in the field of education and of social organization is a new Newton to discover the what and how of an energy of mind which draws minds together and holds them in their orbits of action and thus explains the union of social efforts by which civilization is secured. The distinction in human powers is to be observed in the difference between the action of individual talents and that of an energy of interacting personal character which controls through standards and common ideals and appreciations. Talents are individual abilities acting along some straight road in life. Character, as here used, is an established wisdom and purpose to meet life as a working member in fellowship with all for the common advancement of social welfare. It directs the individual toward that fine ideal described by the word "gentleman."

As indicated at the beginning, research in the educational field of talent training is receiving much attention and a general principle of its inertial action is presented. In the field of character training few definite results have been determined or definite principles formulated. However, the schools and especially the individual teachers have always directed their influence toward maintaining ideals of social responsibility and of establishing definite requirements for such concepts as the care of bodily *Health* to serve a healthy mind; *Loyalty* to associates (one or many) to whom fidelity is due; *Honesty* in conduct and thought; *Cheerfulness* with the spirit of good will; *Courtesy* that gives respect as to an equal and accepts consideration as due from an equal; *Coöperation* with the patience and persistence of a trained morale; *Courage* to face difficulties and dangers with firm determination to carry through; *Industry* that includes determination to accomplish the results sought as a self imposed obligation; and *Self-control* in meeting opposition and in opposing physical or moral obstacles to justified lines of action. To these somewhat fixed directions of energy, teachers have expected their influence to encourage in the able ones an experience of *Leadership* guided by such character elements as those above listed and directed to obtaining group action for worthy ends carried on with the same energizing ideals.

It is evident that an important part of education is involved in this partially recognized and almost entirely unorganized work of individual teachers. To the observer it seems to be a great field for the best research work and one of vital importance.

Reported by An Observer.

G. A. B.

REVIEWS

Address all communications relative to book reviews to Editor Kai Jensen, University of Wisconsin, Madison, Wisconsin.

LACEY, JOY MUCHMORE. *Social Studies Concepts of Children in the First Three Grades.* Contributions to Education, No. 548 (New York: Bureau of Publications, Teachers College, Columbia University, 1932), 89 pp.

This study was undertaken in view of the growing importance of "social studies" in school curricula, the difficulty of curriculum revision and test construction in this field, and the lack of research work on social studies in the elementary grades. The specific problem selected for investigation is "a study of some of the social studies concepts which children in the first three grades have. A sub-problem is to find some evidences of growth in concept formation as shown by the meanings which they associate with them."

The particular concepts to be studied in the main investigation were determined by a preliminary study in which a list of 150 pictures (representing different social studies concepts such as Mother, Easter, Horse, Postman, etc.), and 499 questions about them was made after studying the opinions of educators and classroom teachers regarding social studies curricula, recent representative curricula in the social studies, the content of readers, and personal interviews with 30 children between the ages of six and nine to determine the child's natural interests and activities. The questions and pictures were criticised by college classes in social studies for the primary grades and revised on the basis of this criticism. The use of pictures to represent the concepts is regarded as a valuable and relatively novel feature of this investigation, but it is not clear just how pictures were made for some of the concepts such as Easter, Valentine's day, Summer, Spring, etc.

The reliability of the test was determined by giving it to 100 children and correlating odd with even items (the score on the test being the number of correct answers). The reliability (corrected by the Spearman-Brown formula) of the total test was found to be .809.

For the main study the test was reduced to one hour in length, (125 pictures or concepts, and 315 questions about them). This test was given by the writer or others, to 450 children, divided equally among the first three grades.

Tables are given of the percentage of children in each grade having "misconceptions," defined as cases in which "one or more errors were mingled with the correct responses." These percentages are tabulated separately for each of the 125 concepts, and show, in general, a decrease from the first to the second to the third grade. Tables are also given of the percentage of children in each grade "who have the concepts correct." The meaning of this phrase is not clearly defined, but judging by the results, and the general context it probably means making one or more correct responses. Just why a child should be credited with having a "correct concept" because he has an-

swered one question correctly out of three or four on that concept is a little puzzling. These percentages are also tabulated separately for each of the 125 concepts and show in general an increase from the first to the second to the third grade.

The principal conclusion of the study is that "There seems to be continuous development in children's concepts from grade to grade. The percentage of correct responses steadily increases with progress from grade to grade. At the same time there is a steady decrease in the percentage of misconceptions." Certain inferences are also drawn from the study, such as, there is "need of more definite teaching and testing of the common social studies concepts," "need of an inventory test and a pre-test before the teaching of units of work set forth in the curriculum is attempted," "need of some type of standardized test in social studies to determine whether or not satisfactory results are being secured in the teaching of common concepts of social living."

Certain inferences are drawn from this study which are not justified on the basis of the evidence presented. For instance, it is found in comparing the raw scores on the test for the three grades that the ratio of the difference of the means to the sigma of the difference of the means is larger for grade three minus grade two, than for grade two minus grade one. On the basis of this evidence it is concluded that "there seems to be less difference between the first and second grades than between the second and third grades. The inference is that real differences are not apparent until the beginning of the third grade". The inference might just as well be that there happens to be a large number of questions which are too difficult for the second, but not too difficult for the third grade. Before the evidence presented can be taken to show more difference between the second and third than between the second and first grade one must know that the test in question measures with approximately equal accuracy in different regions of the scale. Since this is not shown, the conclusion is unjustified. Before such conclusions can be affirmed or denied, studies will have to be made much more carefully than this one, taking into consideration recent techniques for evaluating the difficulty of test questions and determining the equality of units at different points on mental measurement scales.

The author recognizes in the earlier part of the work that the test results may be determined not so much by the child's knowledge as by the particular picture and questions used in the test. In general it is recognized that for a response to be "correct" can mean only "correct in so far as this study attempts to determine." Certain of the inferences drawn, however, seem to be in direct violation of this caution. The author does not hesitate to write "From the low percentages of correctness and the various percentages of misconceptions an inference would seem to be that these concepts have not been taught effectively or that they have not been considered of prime importance". Another inference might be that the pictures, the wording of the questions or the scoring of the test are not adequate in some instances. Another conclusion illustrating a similar fallacy is that "concepts which involve personal relationships seem more difficult to teach than mere factual information about objects". Why not infer that such concepts are more difficult to *test?* Specific experiments involving pre-test teaching, testing the results, teaching again, retesting, etc., will have to be made before one can determine whether the diffi-

culty is in the *teaching* or in the *testing* of a given concept.

HAROLD GULLIKSEN
University of Chicago

LULL, HERBERT G. *Secondary Education* (New York: W. W. Norton and Company, Inc., 1932), pp. xvii + 366.

Here is a book that the reviewer really enjoyed reading. In his judgment it is a very sane volume. On every page the author exhibits the virtue of never leaving his reader in doubt as to the position taken in controversial matters—and there are many of them. There are nineteen chapters in the volume. The first five deal with historical matters, and thereby afford a background upon which the rest of the book is projected.

It is our observation that Chapter VI and Chapter XIX constitute two high spots in the author's treatment of his theme. The former deals with "curriculum lag" and the latter with "indoctrination and freedom in education."

Among the factors which contributed to the curriculum lag from 1893 to 1925 are listed the reports of stated committees which perpetuated over-long some faulty pedagogical and psychological assumptions; pride of community in its traditional curriculum; failure of secondary education to parallel life; and so on.

In his treatment of "indoctrination and freedom in education" the author spares no words in denouncing the "child-centrists"; neither does he sympathize with "artificially imposed disciplines." What he pleads for is a middle position in education: "The educational trail lies between the mechanistic view and classical dualism." Although he confesses that he is expressing a faith rather than a demonstrable hypothesis, he feels that without this faith "we are lost in the meshes of institutional determinism."

The author's recommendations in regard to the curriculums of the junior high school, the senior high school, and the junior college exemplify his point of view.

HARLAN C. KOCH
University of Nebraska

GOULD, ADRIAN G., AND DYE, JOSEPH A. *Exercise and Its Physiology.* (New York: A. S. Barnes and Company, 1932), 434 pp.

This volume is prepared with a view to serving "not only as a textbook for students of physical education, and as a reference book for advanced students and teachers of physiology, but also as leisure reading for physicians and others who may be interested in the effects of exercise." The various phenomena which occur in the body during exercise are discussed, and are compared with the occurrences during relative inactivity or rest. In the discussion we find a description of the nervous factors involved in muscle contraction, and of the reflexes which aid in the control of muscle tonus, reciprocal innervation of antagonistic muscles, the reflex and chemical control of respiratory movements and circulatory adjustments. There is also a detailed account of the chemical changes occurring during muscle contraction, fatigue and recovery, including a description of the energy conversions and the oxidase mechanisms. The conveyance of oxygen from the lungs to the tissues is traced, with a discussion of factors governing percentage saturation of the blood with O_2 during rest and exercise, and those governing transfer of oxygen from the capillaries to the tissues. The basic facts involved are organized into a logical presentation of the interrelationships existing between various processes, giving a pic-

ture of adjustments occurring during a short period of exertion and the more permanent changes brought about with longer periods of training.

The presentation of chemical changes is given in considerable detail and a great deal of experimental evidence is given to substantiate the views set forth. In the discussion of the nervous factors involved, however, no mention is made of certain important contributions. For example, the presence of Langley's "receptor substance" between nerve and muscle is accepted. There is no mention of evidence to the contrary, such as that presented by Lovett Evans, which makes very doubtful the authors' statement that "the junctional tissue between muscle and nerve is fatigued sooner than the muscle fiber." The experimental work on chronaxie determinations has shown chronaxie differences between different nerve fibers. The chronaxie of a "synapse" has never been measured, yet it is stated in the text that "in all probability, fatigue products may alter the chronaxie of the synapse or some adjacent structure relative to that of the nerve fiber."

In the field of hormone control, recent evidence by Van Dyke, Bailey and Bucy (1929) has thrown doubt on the actual production of the substance "pituitrin" by the posterior lobe of the hypophysis, making highly speculative Krogh's suggestion given in the text that "it is this hormone which is responsible for the normal capillary tone." In the same way the substances liberated into the blood stream by the medulla and cortex of the adrenal gland have yet to be accurately distinguished and described, so that widespread theorizing as to the presence and effects of adrenalin in normal bodily processes in the intact organism should be only tentatively accepted.

After discussing various nervous and chemical processes, the authors describe interrelationships existing between them, and the changes in these relationships taking place in exercise and training. A few of the factors emphasized by the authors may be briefly set forth. They point to the circulatory system as the weak link in the chain of organs involved in the necessary adjustments of exercise. Its efficiency is primarily dependent upon the efficiency of the heart. The size of the heart is increased through long training and strenuous exercise. The heart of the athlete or trained individual has a slower more powerful beat which accomplishes more work than the more rapid, less forcible beat of the untrained individual. "The functional capacity of the heart determines primarily the limits of a man's power of exertion."

The increase in size of skeletal muscle with training has been shown to be due to a filling out of already existing fibers with sarcoplasm. The increase in strength which occurs in training is thought to be not wholly due to the increase in size of the muscle, but also to a better neuromuscular coördination, more of the formerly little used muscles being brought into play. Neuromuscular coördination is also thought to explain the greater efficiency of an individual to do a certain piece of work in which he has been trained. After training there is elimination of unnecessary and useless muscle responses according to the authors, with elimination of the excess oxygen requirement needed for those muscle responses in the untrained individual. The efficiency of muscles to transform energy into muscular work is not thought to be increased through training. The capacity of the muscles is increased during training, however, through an increase in glycogen content, an increased capacity to tolerate or buffer lactic acid, an increase in phosphocreatine, and possibly an increase in the oxidative agents.

Other topics dealt with in the final chapters of the book are second wind, emotional stress, together with other results of exercise than those mentioned above, completing the picture of complex adjustments occurring during exercise and training.

DOROTHY GULLIKSEN
University of Chicago

TWITMYER, E. B., AND NATHANSON, Y. S. *Correction of Defective Speech.* (Philadelphia: P. Blakiston's Son and Company, Inc., 1932), 413 pp.

The book is labelled by the author as a "manual of psycho-physiological technique for the treatment and correction of the defects of speech." Dr. Twitmyer is professor of psychology and head of the corrective speech clinic at the University of Pennsylvania. Dr. Nathanson is his assistant; and this book is avowedly the outgrowth of this clinical work.

The book begins with the most eloquent foreword that the reviewer ever read as an introduction to a purely professional and technical publication. The authors herein not only delimit their field and view it from a lofty perspective, but also color and outline the details of the field so that they stand out sharply for the reader.

"The whole cosmic scheme of life is applicable to its isolated processes. Speech output — verbal end-result of thought, differential of man and beast, index of genius or fool, characteristic of philosopher or maniac — is one of these isolated processes . . .

". . . *Our attack cannot be successfully waged against an uncontrolled environment; therefore our attack upon the problem must concentrate itself upon the development of the individual.* This philosophic and pragmatic concept has been the basis for our development of the psycho-physiological approach to the diagnosis and correction of the defects of speech—to our recognition of a *minoris locus resistentiae,* an aspect of the *residual diathesis.*"

The foreword is indeed a fine piece of writing and shows that the authors mean to build the structure of their book upon a broad and deep foundation of scholarship. The book intends to include not merely a compilation of recipes but a statement of the rationale of such recipes.

The major part of the book (334 out of 413 pages) is devoted to corrective material and technique. The authors have here included long lists of words and sentences incorporating the sounds upon which the child defective in speech is to be drilled. The authors did not construct these exercises in haphazard fashion, taking each sound and choosing words in which the sound is used without reference to the practical value to the child of the words employed. The authors have chosen 5012 words that appear most commonly in our language. "In computing frequencies and determining classification," they write, "each frequency word was analyzed . . . The purpose of this is apparent. It permits of an analysis of the actual sound frequency properly weighted." Apparently the authors intend not only to give each defective word in the child's vocabulary an amount of drill commensurate with its frequency of use by him, but also to give each defective sound unit as much attention in drill as its frequency in our language would warrant. They include the table of sound frequencies upon which their exercises are built, beginning with (R) at 2,149 occurrences out of 16,668 for all other sounds, and ending with 173 occurrences (CH) out of the same total.

The authors list the sounds that they have thus analyzed in the following pages: xvii, 11, 12, 16, 17, 19. The lists are identical. These lists omit many con-

sonant sounds, among them: the (NG) as in *sing*; (ZH) as in vision, azure, Asia; (Y) as in *y*oung, st(*y*)udent, yet; and its voiceless equivalent, as the initial sound in humorous, human, huge. The omission of these sounds may mean that the authors regard them so infrequent in our language as to need no attention, or that they find these sounds seldom defective.

Under the head of "The Continuants" the authors include only five vowels: (E) as in meet, (AR) as in far, (00) as in moon, (0) as in home, and (I) as in mite. The last two sounds are diphthongs. The authors have omitted at least three diphthongs and five vowels including such infrequent (?) sounds as (A) as in cat, and (OI) as in boy, and (OW) as in cow. (Because the authors have not employed the system of symbolization usually employed in modern words of this kind, *i.e.*, the symbols of the International Phonetic Association, there is some confusion as to what sounds are meant by their letter-symbols. The reviewer here employs the authors' symbolication in order not to be inconsistent.)

At the beginning of each section of exercises the sound to be practiced is described. Many of these descriptions are so vague as to be of no benefit, and in some instances they are misleading. There are many mistakes in the phonetic basis for the practical exercises included in this book. In the discussion of (T), (D) and (N) no mention is made as to which are voiced sounds, while in the analogous relationship among the sounds (P), (B) and (M) is described as possessing "laryngeal vibration" and (B) is said to be the same as (P) "with the exception that the compression is less and there is grosser approximation of the lips," with no mention of the sonant feature of (B). For (N) "the breath is permitted to escape evenly between the oral and nasal passages"—a meaningless statement. With the back plosives (K) and (G) the authors mention that the latter is sonant, but omit to make any mention at all of (NG), which is related to (G) in the same way that (N) is related to (D), and that (M) is related to (B). In discussing the (TH) sounds the authors make no distinctions between the voiced and voiceless sounds, etc., etc. The authors seem to be naive amateurs in the realm of speech sounds.

Since the major contribution of this book is the practical exercises, it leaves the reviewer very little opportunity for commendation. All we can say is that if the teacher knows his phonetics, he will find the word lists helpful; but if he does not know his phonetics, the exercises will be misleading.

ROBERT WEST
University of Wisconsin

JOHNSON, BUFORD J. *Child Psychology*. (Baltimore: Charles C. Thomas, 1932), pp. viii + 439.

In recent years the output of books treating of the child has been out of all proportion to the really significant contributions to knowledge of this period of human development. Writers with one new idea or point of view have either given a distorted picture of the child, or, after completely exploiting their own contribution, have fallen back upon outmoded material or well-worn systems of organization for the remainder of their exposition.

It is safe to say that Dr. Johnson's "Child Psychology" will not be included in the category of "just another text." The material throughout is new. The author's wide experience in original investigations of different developmental levels has enabled her to make a very judicious selection. Old terms rendered

meaningless by use and generalizations so frequently reiterated as to pass for axioms are critically observed and some have new life and precise meaning injected into them.

The writer who would treat of development is faced with a distressing dilemma: shall the organization be along cross-sectional lines or proceed longitudinally? To obtain a genetic picture of individual development the longitudinal effect must be preserved by skillful transition from one cross-section to another. Dr. Johnson has employed a *topical* treatment with longitudinal unfolding within the topic. An organization such as this does not provide a unitary, integrated view of human development from the new-born infant to the adult. The organism as a whole is not seen in the process of development. For this reason it seems apparent that the text was designed not for students of introductory child psychology but rather for advanced students and for intelligent parents with superior education.

Although topical treatment is admirably suited to the needs of the specialist it seems regrettable that the author did not use one consistent system of organization. On the one hand we trace the genetic unfolding of the child in the basic responses which give the chapter headings: Locomotion, Manipulation, and Speech. On the other, we review the characteristics and possibilities of Thought, Perception and Attention, Learning, Growth, Social Behavior and Personality, etc., in the child. To the reviewer it seems that the latter should be made manifest through the evolution of the major activities of the child. Such an unfolding would give us a unified view of individual development and a new and vital type of organization. It would also seem highly essential that the responses serving the general process of alimentation be treated in view of their importance in the life economy of the individual.

The major value of this book, apart from the fresh experimental work incorporated, is contributed by the view, again and again expressed, that the function or role of the educator and parent is to assist in the unfolding of the potentialities of the individual rather than to impress a rigid, conventionalized pattern upon the child; to lead the child through guidance and the introduction of appropriate problems and problem-situations to develop insight into relationships through self-initiated activities rather than to learn imposed material in rote fashion. In the words of the author:

"In the first four years of life the environment of the child should elicit discrimination by use of all the senses. This development comes through specific responses by the child selected on the basis of his experiences . . . Only by self-initiated activities in response to varying forms of stimuli is accuracy of perception developed. The adult may chiefly aid by setting clear patterns of speech and conduct, by providing an environment that offers diversified but not confusing stimulus-patterns, and by giving the child an opportunity to select his responses and to note their effect. (p. 278)

Reasoning develops through solution of problems. Many children are prevented from facing problems through protective habits of adults. A ready-made environment offers no problem for the child to solve. It is often easier and less time-consuming on the part of the child's guardian to solve the problem for the child. Training in rote memory is put over on the child. (p. 338)

Studies of learning show that difficult tasks require exploratory activities for comprehension of the problems and variability in responses before habitual ten-

dencies are established. If we set up standards of leadership, popularity, successful achievement, honesty, politeness, independence, and unselfishness as expectations for child behavior in the first years it is similar to forced guidance in learning. The child may be put through the forms and yet have only a set of stereotyped responses for specific situations without the development of intentions toward social behavior that are flexible in adaptation to varying situations." (p. 365)

A careful reading of this book makes all too apparent one of the major defects of our textbooks in educational psychology, namely, that their orientation is reversed. They give a casual treatment of so-called "original" behavior in the infant and then procede to discuss problems of learning and of motivation on the basis largely of experiments upon adult subjects. They point from the adult infant-wards rather than from the infant adult-wards. Perhaps this explains why our educational system has functioned largely to develop rote learning and our exhortations to "think" have been so barren of results.

For the educated parent Dr. Johnson's book reveals the prospect of an experimental parenthood under which the developing child grows in a material and social environment adequate to his expanding needs, a social environment which provides the best models of speech upon which to pattern his own vocalizations, a world so organized and arranged as to bring about clear-cut and integrated rather than confused responses.

Unquestionably this book represents a major contribution to the fields of educational and child psychology.

KARL C. PRATT
Mount Pleasant State Teachers College

BOGOSLOVSKY, CHRISTINA STAEL VON HOLSTEIN. *The Educational Crisis in Sweden*. (New York: Columbia University Press), pp. xiv + 301.

In the first section of this book the author discusses the national, political, and social background of Swedish education, reviewing the development of Sweden's homogeneous people, its state religion, the nobility, the Crown, and the relationship between the growth of democracy and educational development. Very naturally this discussion embraces both individual and national interests in foreign thought.

The author in the second section of her treatment narrows the discussion to a review of the history of educational activities within the realm, the organization of the school system, various school commissions, and educational bills. In her treatment she outlines the work of the several grades of schools, the various types of educational institutions, private and public, teacher training including traveling scholarships and educational examinations.

The next section is confined to a study of influences of American education on the Swedish system. It includes a survey of periodical literature and accounts and reports of Swedish travel and study in America, together with their effects upon educational thought and practice in Sweden.

The fourth section of this interesting treatise provides a sort of cross section of the educational situation in Sweden in the light of educational experience in America.

In her final chapter the author includes her own concluding remarks and appendices on

1. Advantages of consolidation
2. Student examination and American college entrance examinations

This is a rather unusual study of a foreign school system, reflecting a relatively unconscious imitation of educational procedure in one country by another. Perhaps this type of educational influence is more profound than is effected through a conscious and intentional introduction of foreign procedure. Since educational needs, procedures, and objectives of necessity vary with the locality or with the nation, conscious adaptation too frequently becomes an imposition of procedures admirably suited to one situation but unsatisfactory in another. Improvement in educational systems through comparative studies of educational activities in other lands are profitable only to the degree that they give full account of the racial and national background of thought, traditional attitudes and the aspirations of the people concerned.

This book is well written, interesting, stimulating, and deserves thoughtful reading by the teaching profession.

DAVID ALLEN ANDERSON
Northern Normal and Industrial School

ABELSON, H. H. *The Art of Educational Research: Its Problems and Procedures.* (Yonkers-on-Hudson, New York: World Book Co., 1933), pp. xii + 332.

Within the past eight years at least a dozen general discussions of the methodology of research in education have appeared, exclusive of numerous specialized treatments of the questionnaire, interview, survey, measurements and tests, statistics, experimentation, historical research, case study, legal analysis, activity analysis, speculative thinking, etc. Chronologically arranged, the more general books, intended for use in graduate classes or for field workers, include: Trow, *Scientific Method in Education*, Houghton Mifflin, 1925; Reeder, *How to Write a Thesis*, Public School Publishing Co., 1925 and 1930; Good, *How to Do Research in Education*, Warwick and York, 1928 and 1929; Crawford, *The Technique of Research in Education*, University of Southern California, 1928; Bixler, *Check Lists for Educational Research*, Columbia University, 1928; Alexander, *Educational Research*, Columbia University, 1928, and later revisions; Almack, *Research and Thesis Writing*, Houghton Mifflin, 1930; Waples and Tyler, *Research Methods and Teachers' Problems*, Macmillan, 1930; Whitney, *Methods in Educational Research*, Appleton, 1931; Rusk, *Research in Education*, University of London, 1932; Kelley, *Scientific Method*, Macmillan, 1932; and the volume under review.

The reviewer questions the use of the term "art" in the title of the latest addition to this field, although the author justifies it by defining science as a body of knowledge "formulated with reference to the discovery of general truths or the operation of general laws," and an art as "a systematic application of knowledge and skill in effecting a desired result" (p. 1). He maintains that there may be science *in* education but no science *of* education. It is true that in building up a science of education workers encounter difficulties in determining truth unmistakably and for all time, due to the host of variable factors involved in working with human beings and complex social conditions. For that matter, even workers in mathematics and physics during recent years have not been nearly so complacent as formerly concerning what they once considered a fixed order of things in their respective sciences. In fairness to the author, aside from the title, the reader is not likely to consider educational research a rule-of-thumb application of devices according to established routine rather than a type of critical analysis and fundamental thinking in the search for truth.

The reviewer does not agree with the classification of investigational methodology employed, which is in terms of the purpose for which research is conducted rather than in terms of the basal fact finding procedure involved. This objection may not be serious in view of the probability that varying points of view employed in classifying research are helpful to different types of workers, the present scheme being functional in character. Other writers have classified and discussed research in terms of: fields to which applied, place where conducted, application, data-gathering devices employed, character of the data collected, symbols employed, forms of thinking involved, control of factors, etc. (See "A Symposium on the Classification of Educational Research," *Journal of Educational Research*, May and June, 1931.) As a matter of fact the author, in using a functional classification of research, has included procedures which may be considered as essential in conducting any sound investigation, which possibility he recognizes (p. 47). His five so-called methods of research are: descriptive, causal, evaluative, constructive, and integrative. Would not a comprehensive and adequate study of visual education, for example, include all of the foregoing procedures; a description of present status; causal factors leading to development of this field; evaluation of objectives, materials, or methods involved; application of findings to classroom situations and pertinent recommendations (constructive type); and an analysis of related studies (integrative type).

It is possible that some students may become confused by the elaborateness and nomenclature of some of the subjects, for example, under the descriptive method are socio-historic, socio-survey, genetic-socio, psycho-individual, and psycho-group; and under the causal method are: experimental control, inferred control, comparative, analytical, direct, and transferred. The reviewer's preference is for the older classification (historical, experimental, and normative-survey), based more nearly on research technique than on purpose involved.

In addition to the series of chapters on the different types of research there are discussions of the value of research, selection of a problem for investigation, and writing of the report. Other chapters are devoted to so-called tools of research: statistics, comparison, association, and speculation. Again, these seem to be steps, mental processes, or procedures which might well be employed in any good piece of research. The various chapters are accompanied by "study questions" and well selected references. In addition, a good general bibliography is provided. Aside from the bibliographical references, there are few citations to the literature of research methodology or to available studies as examples of different types of investigation. On the contrary, the book is in the main an attempt to present the author's system of research methodology. As such, it deserves a place in libraries, in graduate classrooms, and in the hands of field research workers.

CARTER V. GOOD
University of Cincinnati

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to A. S. Barr, University of Wisconsin, Madison, Wisconsin.

BULLETINS

A Basis for the Allocation of Funds, Report on Functions of State Institutions of Higher Learning in Mississippi (Nashville, Tennessee: Division of Surveys and Field Studies, George Peabody College for Teachers, 1933), 73 pp.

This bulletin presents a survey of higher education in Mississippi, the present functions of its institutions of higher education, and certain considerations that should be taken into account in determining their functions in the immediate future.

BRAMMELL, P. ROY. *Health Work and Physical Education,* National Survey of Secondary Education, Bulletin No. 17, Monograph No. 28 (Washington: Superintendent of Documents, United States Government Printing Office, 1933), vi + 98 pp.

Bulletin of the National Survey of Secondary Education discussing health and physical education.

BRAMMELL, P. ROY. *Intramural and Interscholastic Athletics,* National Survey of Secondary Education, Bulletin No. 17, Monograph No. 27 (Washington: Superintendent of Documents, United States Government Printing Office, 1933), vi + 143 pp.

Bulletin of the National Survey of Secondary Education discussing intramural and interscholastic athletics.

COOPER, WILLIAM JOHN. *Economy in Education,* School Economy Series Bulletin (Stanford University Press, 1933), 82 pp.

The author presents data upon the need for public school economy, the essential education at the different school levels and suggested economies in public education.

COXE, WARREN W., AND CORNELL, ETHEL L. *Pre-service and In-service Training of Elementary School Teachers:* Suggestions Derived from an Analysis of Teaching Difficulties of Recent Normal School Graduates and from Faculty Criticisms, Educational Monograph of the New York State Teachers Association, No. 2 (Albany, New York: New York State Teachers Association, 1933), xvi + 93 pp.

This bulletin presents an analysis of the teaching difficulties of over 1200 recent normal school graduates teaching in the elementary schools of New York state.

EDDY, HELEN M. *Instruction in Foreign Languages,* National Survey of Secondary Education, Bulletin No. 17, Monograph No. 24 (Washington: Superintendent of Documents, United States Government Printing Office, 1933), vi + 61 pp.

Bulletin of the National Survey of Secondary Education discussing instruction in foreign languages.

EELLS, WALTER CROSBY. *Teachers' Salaries and the Cost of Living*, School Economy Series Bulletin (Stanford University Press, 1933), 94 pp.

The author presents data on the cost of living, and an analysis of the Bureau of Labor Statistics' index number showing its injustice to teachers.

ENGELHARDT, FRED AND OTHERS. *District Organization and Secondary Education*, National Survey of Secondary Education, Bulletin No. 17, Monograph No. 8 (Washington: Superintendent of Documents, United States Government Printing Office, 1933), viii + 208 pp.

A bulletin of secondary education and district organization.

FRANZEN, RAYMOND. *An Evaluation of School Health Procedures*, School Health Research Monographs, No. 5 New York City: American Child Health Association, 450 Seventh Avenue, 1933), xii + 127 pp.

This is the fifth of a series of school health studies published by the American Child Health Association. This study presents factual materials on the elements of a good school health program.

FREEMAN, FRANK N., AND OTHERS. "Pupil Personnel, Guidance, and Counseling," *Review of Educational Research*, III (June, 1933), pp. 185-245.

Contains a summary of research relating to the following topics: school attendance, school progress, school marks, recording and reporting, characteristics of pupil population, guidance and counseling, extra-curricular activities, adjustments and classifications in colleges and universities and special schools and classes. Bibliography.

HAAGE, CATHERINE M. *Tests of Functional Latin for Secondary School Use Based Upon the Recommendations of the Classical Investigation*, Doctor's thesis (Philadelphia, Pennsylvania: University of Pennsylvania, 1932), 192 pp.

This bulletin presents data on the construction and validation of a series of achievement tests in functional secondary school latin.

HAGMAN, ELIZABETH PLEGER. *The Companionships of Preschool Children*, University of Iowa Studies, Studies in Child Welfare, VII, No. 4 (Iowa City, Iowa: University of Iowa, 1933), 69 pp.

From a study of preschool children the author finds that when the measures of individual children in each of nine characteristics were correlated with those of their most frequent companions, higher correlations were obtained than for all pairs. The younger group showed some tendency to choose others similar to themselves in mental age, intelligence quotient, and social reaction, and dissimilar to themselves in extroversion. Chronological age, height, and weight were unrelated. In the older group pairs of most frequent companions tend to be dissimilar in height and weight, and similar in social reaction to the group. Chronological age, mental age, intelligence quotient, extroversion, and a social stimulus index show little or no relationship to the choice of companions.

HATCH, FRIEDA WARDWELL. "Status of the Social Sciences in Secondary Schools of Maine," *The Maine Bulletin*, XXXV, University of Maine Studies, Second Series, No. 27 (Orono, Maine: University of Maine Library, 1933), 114 pp.

The author presents data on the status of instruction in the social studies in the secondary schools of Maine: constants, electives, number of courses offered, size of class, teaching load, the training of the teachers, etc.

KLINE, LINUS WARD, AND CAREY, GERTRUDE L. *A Measuring Scale for Free-Hand Drawing*, Part II, Design and Composition, The Johns Hopkins University Studies in Education, No. 5 (Baltimore: The Johns Hopkins Press, 1933), 58 pp.

The authors present a scale for measuring free-hand drawing with data relative to its standardization.

KRAMER, GRACE A. *The Effect of Certain Factors in the Verbal Arithmetic Problem Upon Children's Success in the Solution*, The Johns Hopkins University Studies in Education, No. 20 (Baltimore: The Johns Hopkins Press, 1933), 106 pp.

This study undertakes to investigate the effects of four factors: interest, sentence form, style-language details, and vocabulary upon the success of sixth-grade children in solving verbal arithmetic problems. While the relationships were uniformly low the highest was found between ability to solve verbal problems and familiarity of vocabulary. The outcome of this study would suggest further careful investigation before schools place too great reliance upon interest content or attractive style to aid to any great extent in developing reasoning power in arithmetic.

LIVESAY, THAYNE M. *A Study of Public Education in Hawaii:* with Special Reference to the Pupil Population, University of Hawaii Research Publications, No. 7 (Honolulu, Hawaii: University of Hawaii, 1932), vii + 120 pp.

This bulletin contains a study of public education in Hawaii: the organization of public education, the population of Hawaii, the general distribution of public school pupils, the race-grade composition of the public school population, and the progress of the public school population. Bibliography.

PRESCOTT, DANIEL A. *The Training of Teachers*, Rutgers University Bulletin, Series IX, Studies in Education, No. 5 (New Brunswick, New Jersey: School of Education, Rutgers University, 1933), 32 pp.

The author presents a criticism of the training of teachers in the United States, and a plan for the training of teachers based upon firsthand contacts with the schools in a number of European countries.

ROBERTS, KATHERINE ELLIOTT. *Learning in Preschool and Orphanage Children*, University of Iowa Studies, Studies in Child Welfare, VII, No. 3 (Iowa City, Iowa: University of Iowa, 1933), 94 pp.

The general purpose of this investigation was to study the process by which young children learn the solution to an initial situation and apply the solution in a subsequent series of similar situations, all of which may be solved according to the same plan. It is shown that both mental and chronological age are important factors in the original solution, but not in the transfer to similar problems.

SCOE, HJALMAR FLETCHER. *Bladder Control in Infancy and Early Childhood*, University of Iowa Studies, Studies in Child Welfare, V, No. 4 (Iowa City,

Iowa: University of Iowa, 1933), 83 pp.

The author presents data on factors conditioning bladder control, the child's responsibility, the expected age of control, sex differences in control, and the features of a practical program of conditioning.

SKEELS, HAROLD MANVILLE. *A Study of Some Factors in Form Board Accomplishments of Preschool Children*, University of Iowa Studies, Studies in Child Welfare, VII, No. 2 (Iowa City, Iowa: University of Iowa, 1933), 148 pp.

This monograph gives a description of a series of form boards for use with preschool children, their validation, and some results obtained with respect to the transfer of training and the effect of conditioning responses.

SMITH, HENRY LESTER, AND NOFFSINGER, FOREST RUBY. *Bibliography of School Buildings, Grounds, and Equipment*, Part II and Part III, Bulletins of the School of Education, Indiana University, IX, Nos. 2 and 3 (Bloomington, Indiana: School of Education, Bureau of Coöperative Research, Indiana University, 1933), 183 and 130 pp.

These bulletins are the second and third volumes of a series of three bulletins giving a bibliography of school buildings, grounds, and equipment. Part I contained references prior to October, 1927; Parts II and III references since October, 1927.

SOPER, WAYNE W. *Development of State Support of Education in New York State*, University of the State of New York Bulletin, No. 1019 (Albany, New York: The University of the State of New York Press, 1933), 66 pp.

This bulletin is the first of a series of studies dealing with educational finance to be published by the Educational Research Division of the New York State Department of Public Instruction. The first bulletin is primarily historical in character.

STANTON, HAZEL M., AND KOERTH, WILHELMINE. *Musical Capacity Measures of Children Repeated After Musical Training*, University of Iowa Studies, Series on Aims and Progress of Research, No. 42 (Iowa City, Iowa: University of Iowa, 1933), 48 pp.

The authors present the results of a ten-year program of testing and retesting by the Seashore Measures of Musical Talent from which they conclude that the Seashore Tests have high scientific value as measures of musical talent.

Theses in Education, Bulletin of the Bureau of School Service, V, No. 4 (Lexington, Kentucky: College of Education, University of Kentucky, 1933), 56 pp.

An annotated bibliography and abstract of theses in education, prepared at the University of Kentucky, from the initiation of graduate work to January 1, 1933.

WATSON, GLADYS HIPPLE. *Success and Failure in the Teaching Profession*, Doctor's Dissertation (New York: Department of Psychology, Teachers College, Columbia University, 1932), 99 pp.

The author presents the results of a number of personality and vocational tests for a group of 58 students in Teachers College, Columbia University. The author brings the items from the several tests having predictive value together into a single representative scale.

ZEIGEL, WILLIAM H. *Research in Secondary Schools*, National Survey of Secondary Education, Bulletin No. 17, Monograph No. 15 (Washington: Superintendent of Documents, United States Government Printing Office, 1933), vi + 72 pp.

This bulletin of the National Survey of Secondary Education discusses research in secondary schools.

EDUCATIONAL ADMINISTRATION AND SUPERVISION

ACKLEY, C. E. "Constitutional Control of the Curriculum," *Educational Law and Administration*, II (July, 1933), pp. 64-66.

The author presents the results of an investigation of constitutional restrictions on legislation regarding the curriculum.

CHAMBERS, M. M. "Public Aid for Colleges Under Private Control," *Educational Law and Administration*, II (July, 1933), pp. 53-63.

The author presents a digest of legislation relating to public aid for privately controlled colleges and universities.

HYDE, ANNA, AND HYDE, RICHARD. "Vocational Misplacements Among Elementary-School Teachers," *Elementary School Journal*, XXXIII (June, 1933), pp. 746-752.

From a study of the placement of 3955 elementary teachers the authors show that regardless of the size of the city there is a tendency among school officials in the placement of teachers to ignore the teacher's special field of training.

THOMPSON, CLEM O. "The Life-Certificate," *Elementary School Journal*, XXXIII (June, 1933), pp. 762-768.

The author presents certain data relative to the life-certification of teachers and raises a question concerning the worthwhileness of the policy.

EXPERIMENTAL STUDIES OF LEARNING AND TEACHING

BRUCE, R. W. "Conditions of Transfer of Training," *Journal of Experimental Psychology*, XVI (June, 1933), pp. 343-361.

This study is an investigation of some of the conditions of transfer of training. The results are formulated in a set of laws. The article contains numerous references to previous investigations.

COX, JOHN W. "Some Experiments on Formal Training in the Acquisition of Skill," *British Journal of Psychology*, XXIV (July, 1933), pp. 67-87.

An experimental study of the transfer of training in the acquisition of skill where two methods of instruction, the "practice" method and the "training" method, are used in which the author finds that mere repetition confers little advantage upon the performance of other operations, but that where repetition is replaced by suitable instruction the skill developed at no additional cost of time tends to transfer to other operations over a wide field of manual activities.

GARRISON, K. C. "A Study of Some Factors Related to Leadership in High School," *Peabody Journal of Education*, XI (July, 1933), pp. 11-17.

After reviewing previous investigations the author presents the results of a statistical study of the relationship between school average, height, weight, father's occupational rating, number in family, position in family, and leadership. The correlations were uniformly

low, the highest (.36) being with the school average.

GUILER, WALTER SCRIBNER. "Improvement and Permanency of Learning Resulting from Remedial Instruction," *School Review*, XLI (June, 1933), pp. 450-458.

An experimental study in which the author finds that the progress in the remedial group was more marked than the non-remedial group.

JENKINS, J. G. "Instruction as a Factor in 'Incidental' Learning," *American Journal of Psychology*, XLV (July, 1933), pp. 471-477.

After reviewing previous investigations the author reports the results of an experimental study from which he concluded that "intentional" learning is superior to "incidental" learning.

LITTERER, OSCAR. "An Experimental Study of Visual Apprehension in Reading, *Journal of Applied Psychology*, XVII (June, 1933), pp. 266-276.

Through the application of a number of tests the author attempts to discover the true relationship between visual apprehension and achievement in reading.

OLANDER, HERBERT T. "The Need for Diagnostic Testing," *Elementary School Journal*, XXXIII (June, 1933), pp. 736-745.

The author presents data to show that while the forty teachers involved in this study were able to diagnose the exact causes of a pupil's errors in multiplication with a high degree of accuracy, their success in diagnosing errors in addition, subtraction, and division was limited. The average of all the teachers was approximately 30 percent or less than one error out of three correctly diagnosed.

PICKFORD, R. W. "The Tendency Towards Synthesis in Reading," *British Journal of Psychology*, XXIV (July, 1933), pp. 50-66.

An experimental study of reading in which the author finds a tendency towards synthesis in the reading of simple and compound passages, each passage being treated as a unit only when the subject completely fails to unify its parts. The tendency to unification is fundamental, giving to the materials real meaning.

SYMONDS, PERCIVAL M. "The Psychology of Problem-Solving," *Harvard Teachers Record*, III (June, 1933), pp. 137-143.

The author makes a plea for more information about the mental processes involved in problem-solving and the importance of problem-solving.

CURRICULUM CONSTRUCTION

CHARTERS, W. W., AND HINDMAN, DARWIN A. *The Duties of Ohio Public-Health Commissioners*, Ohio State University Studies, Bureau of Educational Research Monographs, No. 17 (Columbus, Ohio: Ohio State University, 1933), xi + 70 pp.

This bulletin presents an analysis of the duties of Ohio Public-Health Commissioners for the purpose of curriculum construction. The bulletin contains a statement of the procedure, the activities of Public-Health Commissioners in Ohio, and suggestions on constructing courses of training.

CLEM, ORLIE M., AND DODGE, SEWARD B. "The Relation of High-School Leadership and Scholarship to Post-School Success," *Peabody Journal of Education*, X (May, 1933), pp. 321-329.

A study of the relationship of high-school leadership to post-school success in which the author concludes that those graduates who were leaders in high school have made a better showing in post-school activities than have scholars or members of the random group.

KEFAUVER, GRAYSON N., AND MACKENZIE, GORDON N. "The Value of Secondary School Subjects in the Preparation for Law," *California Quarterly of Secondary Education*, VIII (June, 1933), pp. 383-394.

The authors present an analysis of the secondary-school subjects required in the preparation for law, and the importance of these subjects as judged by deans of law schools.

RANDALL, JOHN A., CHARTERS, W. W., AND ELLINGSON. "Education and Research at a Mechanics Institute: I. Reconciling Technical, Liberal, and Research Objectives; II. Selecting Institutional Objectives; III. Activity Analysis as a Basis for Course Content," *Personnel Journal*, XIII (June, 1933), pp. 1-15.

The three articles explain the steps taken during the last several years to ascertain the needs of their educational constituency and to reconstruct curriculum, content, techniques, and administrative practices in harmony with re-defined aims and with specific objectives whose feasibility has been experimentally ascertained.

REAVIS, WILLIAM C. "Interscholastic Non-athletic Activities in Selected Secondary Schools," *School Review*, XLI (June, 1933), pp. 417-428.

From an analysis of a questionnaire from which 224 replies were received, the author finds that there are many opportunities for pupil practice and participation in a great variety of interscholastic non-athletic activities in secondary schools.

STARBIRD, MYRTLE, AND OTHERS. "Out-of-School Uses of English," *The English Journal*, XXII (June, 1933), pp. 466-471.

The author presents an analysis of the out-of-school uses of English by 345 high-school pupils. Tables and index numbers.

WILLIAMS, L. A. "Analysis Techniques in Curriculum-making," *School Review*, XLI (June, 1933), pp. 437-442.

The author presents an analysis of the job analysis, activity analysis, functional analysis, and content analysis techniques in curriculum construction.

MEASUREMENTS, STATISTICS, AND METHODOLOGY OF SCIENTIFIC RESEARCH

GREENE, HARRY A., AND BETTS, EMMETT A. "A New Technique for the Study of Oral-Language Activities," *Elementary School Journal*, XXXIII (June, 1933), pp. 753-761.

The authors present data on the percentage of errors made by ordinary recording devices and a description of a new language recording apparatus. The apparatus may be used for a variety of recording purposes.

HOLZINGER, KARL J., AND SWINEFORD, FRANCES. "Selected References on Statistics and Theory of Test Construction," *School Review*, XLI (June 1933), pp. 462-466.

An annotated bibliography on statistics and the theory or test construction of thirty selected titles.

RESEARCH NEWS and COMMUNICATIONS

Address all research news and communications to Carter V. Good, Teachers College, University of Cincinnati, Cincinnati, Ohio.

SELECTED BIBLIOGRAPHY ON TECHNIQUES OF EDUCATIONAL RESEARCH AND RELATED PROBLEMS, 1932-33[1]

CARTER V. GOOD
University of Cincinnati

THIS, the fourth annual bibliography on research techniques to be published in the *Journal of Educational Research*, is intended to be of service to three groups of individuals: field workers interested in research, graduate students, and staff members of higher institutions.

The bibliography covers the period from approximately June 1, 1932, to July 1, 1933. Limitations of space have made it necessary to select those references of greatest practical value to field workers, graduate students, and staff members of teacher-training institutions. Therefore, highly technical discussions of a statistical and mathematical character have been omitted.

The centers of interest for the past year, as represented in the bibliography, are: evaluations and predictions of progress in given fields; an increasing number of bibliographies, summaries of educational literature, and abstracts of graduate theses; needed studies; methodology of research; and training and guidance of graduate students and research workers.

[1] Supplementary to bibliographies previously published in the *Journal of Educational Research*, October, 1930; September, 1931; October, 1932; and Carter V. Good, *How To Do Research in Education*. Baltimore: Warwick and York, 1929.

CHARACTERIZATION AND EVALUATION OF RESEARCH AND EDUCATIONAL PROCEDURE

Almack, J. C. "Scientific Method in Teaching," *Educational Method*, XII (March, 1933), 321-24.

Barnes, H. E. "The Development of Sociology," *Scientific Monthly*, XXXV (December, 1932), 543-53.

Bernard, L. L., and J. S. "A Century of Progress in the Social Sciences," *Social Forces*, XI (May, 1933), 488-505.

Briggs, T. H. "A Vision of Secondary Education," *Teachers College Record*, XXXIV (October, 1932), 1-17.

Caldwell, O. W. "Science—Truth or Propaganda," *School and Society*, XXXVII (February 11, 1933), 169-74.

Charters, W. W. "The Value of a Research Organization," *Educational Research Bulletin*, XI (October 26, 1932), 351-52.

Cox, P. W. L. "The Research Racket," *Junior-Senior High School Clearing House*, VII (January, 1933), 260-65. A criticism of questionable practices in educational investigation.

Davis, R. A. "Research and the Schools," *Journal of Educational Research*, XXVI (April, 1933), 561-68.

Douglass, H. R. "Does the Future of Secondary Education Depend upon Statistical Research in Foundational

Thinking?'' *Education*, LIII (January, 1933), 284-89.

Eginton, D. P. ''Research, the Fourth 'R','' *Journal of Education*, CXV (October 17, 1932), 574-75.

Gee, W. P. *Research Barriers in the South*. New York: Century Co., 1932. Pp. 192.

Gray, J. S. ''Scientific Postulates for Educational Research,'' *Educational Administration and Supervision*, XIX (January, 1933), 18-24.

Harap, Henry, and Bayne, Alice J. ''A Critical Survey of Public School Courses of Study Published 1929 to 1931,'' *Journal of Educational Research*, XXVI (October, 1932), 105-09.

Howell, C. E. ''The Functions of a Research Department,'' *American School Board Journal*, LXXXV (November, 1932), 50, 64.

Hutchins, Robert M. ''The American Educational System,'' *School Review*, XLI (February, 1933), 95-100.

Judd, C. H. ''The Scientific Development of the Curriculum,'' *Elementary School Journal*, XXXIII (April, 1933), 577-91.

Judd, C. H. ''In Defense of American Secondary Schools,'' *School Review*, XLI (April, 1933), 257-65.

Lambert, A. C. ''The Research Point of View in Education,'' *American School Board Journal*, LXXXVI, (March, 1933), 16.

Lindquist, E. F., and Cook, Walter W. ''Experimental Procedures in Test Evaluation,'' *Journal of Experimental Education*, I (March, 1933), 163-85.

Myers, Alonzo F. ''An Evaluation of Research on Teacher Preparation,'' *Journal of Educational Research*, XXVI (November, 1932), 161-74.

Odell, C. W. ''Educational Measurement in the Secondary School,'' *Journal of Educational Research*, XXVI (October, 1932), 81-89.

Reymert, M. L. ''An Attempt at a Rating Scale for Psychologists,'' *Psychological Bulletin*, XXX (January, 1933), 63-79.

Sackett, E. B. ''Is Research an Expense?'' *American School Board Journal*, LXXXVI (April, 1933), 19-20.

Touton, F. C. ''Education in 1940,'' *Education*, LIII (March, 1933), 428-35.

Woody, Clifford. ''Whither Educational Research?'' *Journal of Educational Research*, XXVI (October, 1932), 136-37.

Problems and Issues: Needed Research

Capen, S. P. ''Report of the Committee on Problems and Plans in Education,'' *Educational Record*, XIII (July, 1932), 180-84.

Carr, W. G. ''New Angle of Attack Needed in Class Size Research,'' *Nation's Schools*, X (November, 1932), 27-30.

Chambers, M. M. ''Needed Studies in Educational Administration,'' *Educational Administration and Supervision*, XIX (April, 1933), 268-74.

Coon, Beulah I., and others. *Suggestions for Studies and Research in Home-Education: Outline of Projects Needing Investigation and Some Basic Considerations in Relation to Them*. Bulletin No. 166. Washington: Federal Board for Vocational Education, 1932. Pp. xi + 77.

Cunningham, Bess V., and Connor, Ruth. ''Suggested Experiments in the Field of Pre-parental Education,'' *Teachers College Record*, XXXIV (January, 1933), 285-301.

''Current Research Projects (1932),'' *American Journal of Sociology*, XXXVIII (March, 1933), 728-54.

Good, Carter V. "Doctors' Theses Under Way in Education, 1932-1933," *Journal of Educational Research*, XXVI (January, 1933), 381-400.

Lee, J. M. "Needed Developments in Measurement for Secondary Schools," *Education*, LIII (March, 1933), 424-27.

McCloy, C. H. "Research in Athletics for Girls and Women," *Research Quarterly of the American Physical Education Association*, III (October, 1932), 101-04.

New York State Education Department, *Needed Educational Research in New York State*. Albany: University of the State of New York Press, 1933. Pp. 16.

"Report of Researches in Education," *British Journal of Educational Psychology*, II (1932), 221-27. A list of psychological studies under way in Great Britain, which are of interest to workers in education.

Symonds, P. M. "Needed Research in the Teaching of English," *English Journal, College Edition*, XXII (June, 1933), 447-56.

Taylor, A. S. "Curricular Research Is Urgently Needed," *Junior College Journal*, III (February, 1933), 246-48. Deals with junior college curricular problems.

Touton, F. C. "Research Projects of the Secondary School Level Carried on in California Cities During 1931-1932," *California Quarterly of Secondary Education*, VIII (January, 1933), 189-219.

SOURCES OF INFORMATION: BIBLIOGRAPHIES, SUMMARIES, AND ABSTRACTS

Abel, J. F. "Selected References on Foreign Education," *Elementary School Journal*, XXXIII (June, 1933), 777-84.

"Abstracts of Unpublished Masters' Theses, Indiana State Teachers College, 1931 and 1932," *Teachers College Journal*, III (July, 1932), 251-300.

"Abstracts of Unpublished Masters' Theses, Indiana State Teachers College, 1933," *Teachers College Journal*, IV (July, 1933), 277-320.

Affleck, G. B. "Bibliography for 1931," *Research Quarterly of the American Physical Education Association*, III (December, 1932), 122-54. Covers the field of physical education.

Alexander, Carter. "Selecting Magazines for Teachers Starved for Professional Vitamins," *School Executives Magazine*, LII (April, 1933), 268-70.

An Analytical Bibliography of Modern Language Teaching, 1927-1932. Compiled for the Committee of Modern Language Teaching by Algernon Coleman, with the assistance of Agnes Jacques. Chicago: University of Chicago Press, 1933. Pp. xiv + 296.

Angrist, Charles K. "The Technique of Library Research for Students of Education." Unpublished Master's Thesis, College of the City of New York, 1931.

Baker, H. J., and others. "Tests of Intelligence and Aptitude," *Review of Educational Research*, II (October, 1932), 271-342. Includes 450 references.

Blauch, L. E. "Curriculum Surveys in Higher Education," *Journal of Higher Education*, IV (May, 1933), 255-60. Reviews nineteen curriculum surveys.

Brownell, William A. "The Growth and Nature of Research Interest in Arithmetic and Reading," *Journal of Educational Research*, XXVI (February, 1933), 429-41.

Campbell, Doak S. "Directory of the Junior College, 1933," *Junior College Journal*, III (January, 1933), 217-31.

Cattell, J. M. "The Distribution of American Men of Science in 1932," *Science*, LXXVII (March 10, 1933), 264-70.

Cook, P. A. W. "A Guide to the Literature on Negro Education," *Teachers College Record*, XXXIV (May, 1933), 671-77.

Crabb, A. L., and others. "Bibliography of Studies Dealing Wholly or in Part with the Training School," *Peabody Journal of Education*, X (September, 1932), 98-114.

"Cubberley (A Series of Articles on His Life and Work)," *California Quarterly of Secondary Education*, VIII (April, 1933), 229-49.

"Educational Books of 1932," *School and Society*, XXXVII (March 25, 1933), 391-408.

Eells, W. C. "1933 as a Centennial Year in the History of Education," *School and Society*, XXXVII (January 7, 1933), 12-13.

Eliassen, R. H., and Anderson, E. W. "Investigation of Teacher Supply and Demand Reported Since November, 1931," *Educational Research Bulletin*, XII (March 8, 1933), 66-72.

Fink, Stuart D., and Umstattd, J. G. *Abstracts of Masters' and Doctors' Theses in Education, University of Minnesota, July 1, 1930, to July 1, 1931.* Educational Research Bulletin No. 4, Eta Chapter, Phi Delta Kappa, College of Education, University of Minnesota. Minneapolis: Phi Delta Kappa, University of Minnesota, 1932. Pp. 72.

Freeman, F. N., and others. "Selected References on Educational Psychology, *School Review*, XLI (May, 1933), 379-86.

Friedman, Kopple C., and Nemzek, Claude L. "The School Magazine in Educational Literature," *School Review*, XL (October, 1932), 620-26.

Good, Carter V. "Selected Bibliography on Techniques of Educational Research and Related Problems, 1931-1932," *Journal of Educational Research*, XXVI (October, 1932), 148-60.

Gray, William S. "Summary of Reading Investigations (July 1, 1931, to June 30, 1932)," *Journal of Educational Research*, XXVI (February, 1933), 401-24.

Heck, A. O., and others. "Pupil Personnel, Guidance, and Counseling," *Review of Educational Research*, III (June, 1933), 183-278. Includes 793 references.

Hildreth, Gertrude H. *Bibliography of Mental Tests and Rating Scales.* New York: Psychological Corporation, 1933.

Hollingworth, L. S. "Selected References from the Literature on Exceptional Children," *Elementary School Journal*, XXXIII (May, 1933), 698-706.

Holy, T. C., and others. "School Buildings, Grounds, Equipment, Apparatus, and Supplies," *Review of Educational Research*, II (December, 1932), 343-442. Includes 476 references.

Holzinger, K. J., and Swineford, Frances. "Selected References on Statistics and Theory of Test Construction," *School Review*, XLI (June, 1933), 462-66.

Joyal, A. E. *Abstracts of Doctors' Theses in Education at the University of California, 1898 to March, 1933.* Berkeley, California: Lambda Chapter, Phi Delta Kappa, 1933. Pp. 104.

Kitson, H. D., and Stoyer, E. M. "Measuring Vocational Guidance: A Summary of Attempts," *Personnel Journal*, XI (October, 1932), 150-59.

Koos, Leonard V., and others. "Selected References on Secondary-School Instruction," *School Review*, XLI (January, February, March, 1933), 60-66, 129-46, 213-26.

Lancaster, J. H. "A Guide to the Literature on Education of Teachers," *Educational Administration and Supervision*, XIX (May, 1933), 363-72.

Lee, J. M., and Symonds, P. M. "New Type or Objective Tests: A Summary of Recent Investigations," *Journal of Educational Psychology*, XXIV (January, 1933), 21-38.

Louttit, C. M. *A Handbook of Psychological Literature.* Bloomington, Indiana: Principia Press, 1932.

McCloy, C. H. "Bibliography of Health and Physical Education," *Research Quarterly of the American Physical Education Association*, III (October, 1932), 3-78.

McGeoch, J. A. "The Psychology of Human Learning: A Bibliography," *Psychological Bulletin*, XXX (January, 1933), 1-62.

McLaughlin, Katherine L. "Selected References on Kindergarten-Primary Education," *Elementary School Journal*, XXXIII (April, 1933), 621-26.

Meister, Morris. "Recent Educational Research in Science Teaching," *School Science and Mathematics*, XXXII (November, 1932), 875-89.

Mikesell, Homer Earl. "The Educational Research Activities of the State of Ohio." Unpublished Master's thesis, Ohio State University, 1930.

Peters, C. C., and Struck, F. T. *Abstracts of Studies in Education at the Pennsylvania State College, Part II.* Pennsylvania State Studies in Education, No. 4. State College, Pennsylvania, 1932. Pp. 64.

Pieper, C. J. "Research Studies Relating to the Teaching of Science," *Science Education*, XVII (April 1933), 138-50. Covers the period from June, 1931, to December, 1932.

Reavis, W. C., and Henry, N. B. "Selected References on Public School Administration," *Elementary School Journal*, XXXIII (January, February, 1933), 382-86, 462-66.

Roemer, Joseph. *Abstracts of Unpublished Masters' Theses in the Field of Secondary-School Administration.* Bulletin of the Department of Secondary-School Principals of the National Education Association, No. 43, 1932. Cicero, Illinois: H. V. Church, Secretary.

Selle, Erwin S. *The Organization and Activities of the National Education Association.* Teachers College Contributions to Education, No. 513. New York: Bureau of Publications, Teachers College, Columbia University, 1932.

Shannon, J. R. "An Analysis of the Subject Matter of Education Discussed in Lay Magazines," *School and Society*, XXXVII (June 3, 1933), 715-16.

Shannon, J. R. "Professional Experience of Professors of Education," *School and Society*, XXXVI (November 12, 1932), 638-40. Based on data from *Leaders in Education.*

"Sixty Educational Books of 1932," *Journal of National Education Association*, XXII (April, 1933), 117-20.

Smith, H. L., and Noffsinger, F. R. *Bibliography of School Buildings, Grounds, and Equipment.* Bulletin of the School of Education, Indiana University, Vol. IX, No. 2, March, 1933. Bloomington, Indiana: Indiana University, 1933. Pp. 182.

Stenquist, J. L., and others. "Educational Tests and Their Uses," *Review of Educational Research*, III (February, 1933), 1-80. Includes a bibliography of 467 items.

Stint, D. B., and Engelhart, M. D. "A Critical Summary of the Research on the Lecture-Demonstration Versus the Individual-Laboratory Method of Teaching High School Chemistry," *Science Education*, XVI (October, 1932), 380-91.

Stoddard, G. D., and others. "Mental and Physical Development," *Review of Educational Research*, III (April, 1933), 81-181. Bibliography of 433 items.

Terry, Paul W. "Summary of Investigations of Extra-Curriculum Activities in 1931, I," *School Review*, (September, October, 1932), 505-14, 613-19.

Terry, Paul W. "Selected References on the Extra-Curriculum," *School Review*, XLI (April, 1933), 299-306.

Theses in Education, University of Kentucky. Bulletin of the Bureau of School Service, Vol. V, No. 4, June, 1933. Lexington, Kentucky: University of Kentucky, 1933. Pp. 62.

Tinker, M. A. "Diagnostic and Remedial Reading," *Elementary School Journal*, XXXIII (December, 1932, and January, 1933), 293-306, 346-57. A summary of studies in this field.

Witmer, Eleanor M., and Miller, Margaret C. "United States Office of Education Serial Publications," *Teachers College Record*, XXXIV (January, 1933), 302-11.

Wright, Edith A. *Bibliography of Research Studies in Education: 1930-1931.* Office of Education Bulletin, No. 16, 1932. Washington: Office of Education. Pp. xviii + 460.

Wright, Edith A. *List of Educational Research Studies of State Departments of Education and State Education Associations, 1931-1932.* Circular No. 63, October, 1932. Washington: Office of Education. Pp. 44.

Young, Grace P. "Bibliography of Modern Language Methodology in America for 1932," *Modern Language Journal*, XVII (May, 1933), 600-615.

Collecting, Analyzing, Interpreting, and Reporting Data

Abelson, Harold H. *The Art of Educational Research: Its Problems and Procedures.* Yonkers-on-Hudson, New York: World Book Co., 1933. Pp. xii + 332.

Barr, A. S. "The Integration of Scientific Techniques in Educational Research," *Journal of Educational Research*, XXVI (November, 1932), 219-21.

Breed, Frederick S. "The Nature of Educational Measurement," *School Review*, XL (September, 1932), 497-504.

Campbell, L. R. "Editing the Curriculum," *School and Society*, XXXVII (February 25, 1933), 251-56.

Caswell, H. L. "Definition and Classification of Surveys," *Peabody Journal of Education*, X (November, 1932), 179-81.

Cavan, Ruth S. "The Questionnaire in a Sociological Research Project," *American Journal of Sociology*, XXXVIII (March, 1933), 721-27.

Cozens, F. W. "The Technique of a School Survey in Health and Physical Education," *Research Quarterly of the American Physical Education Association*, III (December, 1931), 3-20.

Davis, E. C. "The Survey and Professional Preparation in Physical Education," *Research Quarterly of the American Physical Education Association*, III (December, 1932), 21-32.

Davis, R. A., and Wilbur, M. J. "Curriculum Research," *Journal of Experimental Education*, I (June, 1933), 294-97. An analysis of purposes, methods, and devices of research in this field from 1921-31.

Edwards, Newton. *The Courts and the Public Schools.* Chicago: University of Chicago Press, 1933. Pp. xvi + 582.

Ellwood, C. A. "Scientific Method in Sociology," *Social Forces*, XI (October, 1932), 44-50.

Foster, F. M. "Uniformity in Adminis-

trative Nomenclature," *Junior College Journal*, III (April, 1933), 362-64.

Kelley, T. L. *Scientific Method:* Its Function in Research and in Education. New York: Macmillan Co., 1932. Pp. x + 234.

Knudsen, Charles W. "Critical Study with a Minimum of Research," *Journal of Educational Research*, XXVI (February, 1933), 457.

Monroe, Walter S. "Standardization of Statistical Symbolism," *Journal of Experimental Education*, I (March, 1933), 223-28.

Monroe, Walter S., and Stuit, Dewey B. "The Interpretation of the Coefficient of Correlation," *Journal of Experimental Education*, I (March, 1933), 186-203.

Morrow, P. R., and Mishoff, W. O. *A Guide to Thesis Writing.* Athens, Georgia: University of Georgia, 1932. Pp. 20.

Rusk, Robert R. *Research in Education:* An Introduction. London, England: University of London Press, Ltd., 1932. Pp. 108.

Sears, J. B. "The Function of Research in Curriculum Making," *California Quarterly of Secondary Education*, VIII (January, 1933), 147-53.

Smith, Mapheus. "A Note on Stability in Questionnaire Response," *American Journal of Sociology*, XXXVIII (March, 1933), 713-20.

Toops, Herbert A. "Some Considerations Relative to the Standardization of Certain Procedures in Educational Research," *Journal of Experimental Education*, I (March, 1933), 229-38.

Upshall, C. C., and Masters, H. V. "Research in Education — Interpretive Hazards," *School and Society*, XXXVI (December 31, 1932), 851-53.

West, Paul V. "Need for Standardization of Symbols and Formulae in Educational Statistics," *Journal of Experimental Education*, I (March, 1933), 216-22.

Whitley, Robert Love. "The Case Study as a Method of Research." Unpublished Doctor's thesis, New York University, 1931.

Woodard, J. W. "Critical Notes on the Nature of Sociology as a Science," *Social Forces*, XI (October, 1932), 28-43.

Woody, Clifford, and Sangren, Paul V. *Administration of the Testing Program.* Yonkers-on-Hudson, New York: World Book Co., 1932. Pp. xii + 398.

Worcester, D. A. "Has Standardized Testing Been Over-emphasized?" *School Executives Magazine*, LI (July, 1932), 467-69, 490.

Training Research Workers and Guidance of Graduate Students

Agunod, Nemesio L. "The Bureau of Research and Reference in American School Systems." Doctor's thesis, Teachers College, Columbia University, 1931.

Betts, G. H. "General Information Possessed by Graduate Students in Education," *School and Society*, XXXVII (June 24, 1933), 821-24.

Betts, G. H. "Subsidizing Graduate Study," *Journal of Higher Education*, III (November, 1932), 415-18.

Brueckner, L. J. "Planned Programs of Research," *Journal of Educational Research*, XXVI (January, 1933), 364-66.

Carpenter, W. W., and Rosenstengel, W. E. "Participation of State Universities in School Plant Surveys," *School Executives Magazine*, LII (November, 1932), 122-23.

Clark, Robert. "The Mobility of Public School Administrators," *School and Society*, XXXVI (October 15, 1932), 506-08.

Charters, W. W., and others. "Annual

Report of the Bureau of Educational Research, Ohio State University," *Educational Research Bulletin*, XI (October 12, 1932), 309-36.

Cooper, W. J. "The Scientific Work of the Government of the United States," *Scientific Monthly*, XXXVI (February, 1933), 121-30.

Cooper, W. J. "Professional Leadership in Education," *School and Society*, XXXVII (March 11, 1933), 305-12.

Coxe, W. W. "The Coördination of Research by State Departments of Education," *School and Society*, XXXVII (June 10, 1933), 729-35.

Edmonson, J. B. "A University's Program of Research in Elementary Education," *Elementary School Journal*, XXXIII (October, 1932), 106-11.

Fernberger, S. W. "Objective Control of One Factor in the Granting of Post Graduate Fellowships and Scholarships," *School and Society*, XXXVII (June 17, 1933), 787-92.

Gilbert, L. C. "Professional Reading of Young Students of Education," *School Review*, XL (October, 1932), 606-12.

Good, Carter V. "Cautions to the Beginner in Educational Research," *Journal of Educational Research*, XXVI (December, 1932), 302-4.

Harris, R. G. "The Doctorate and the Depression," *School and Society*, XXXVI (October 15, 1932), 498-500.

Hull, Callie, and West, C. J. "The Eighth Census of Graduate Research Students in Chemistry, 1931," *Journal of Chemical Education*, IX (August, 1932), 1472-73.

Hull, Callie, and West, C. J. "The Ninth Census of Graduate Research Students in Chemistry, 1932," *Journal of Chemical Education*, X (August, 1933), 499-503.

Larson, E. L. "Major Problems Relating to the Master's Degree in Education," *Educational Administration and Supervision*, XVIII (December, 1932), 699-711.

MacNeel, Joseph Raymond. *Admission of Students as Candidates for Master's Degree.* Teachers College Contributions to Education, No. 524. New York: Bureau of Publications, Teachers College, Columbia University, 1932.

Monroe, W. S. "Doctors' Degrees in Education, 1928-1933," *School and Society*, XXXVII (March 4, 1933), 289-90.

Nelson, R. W. "Theaetetus Ph.D.," *Journal of Higher Education*, IV (May, 1933), 234-40. Criticism of restrictions placed upon Doctors' candidates in working out their programs and theses.

Rogers, L. B. "The Place of Schools of Education in the Program of Improvement of University Instruction," *School and Society*, XXXVI (August 27, 1932), 257-64.

Ruckmick, C. A. "The Future of the Doctoral Degree," *School and Society*, XXXVII (February 18, 1933), 201-6.

Scates, D. E. "Training for Research in the Light of Field Requirements," *Elementary School Journal*, XXXIII (October, 1932), 96-105.

Sears, Jesse B. "The Administration of Public School Research Policies," *Journal of Educational Research*, XXVI (November, 1932), 186-98.

Shannon, J. R. "Geographical Distribution of Teachers in Education," *School and Society*, XXXVI (December 3, 1932), 720-22. Based on an analysis of "Leaders in Education: A Biographical Directory."

Strang, Ruth. "Vocational Paths to Certain Educational Positions," *School and Society*, XXXVI (October 15, 1932), 508-12.

Stratton, Dorothy C. *Problems of Students in a Graduate School of Educa-*

tion. Teachers College Contributions to Education, No. 550. New York: Teachers College, Columbia University, 1933.

Symonds, P. M. "Characteristics of a Good Dissertation in Education," *Teachers College Record,* XXXIV (January, 1933), 312-26.

Toops, Laura C. "Research for Teachers in Training," *Teachers College Record,* XXXIV (December, 1932), 228-34. A plan for coördinating coöperative and individual research applied in a research program for a class of prospective teachers in nursery schools.

"Training for Research in the Social Sciences," *Educational Record,* XIV April, 1933), 152-61. Discusses requirements for graduate degrees.

Walker, Helen M. "Problems in the Training of Research Workers," *Journal of Educational Research,* XXVI (February, 1933), 442-48.

West, Clarence J., and Hull, Callie. "Doctorates in Chemistry and Related Fields Conferred by American Universities, 1931-32," *Journal of Chemical Education,* IX (November, 1932), 1975-96.

Zeigel, William H., Jr. "Research for Secondary Schools," *School Review,* XLI (January, 1933), 27-39.

Zeigel, William H., Jr. *Research in Secondary Schools.* Office of Education Bulletin, No. 17, 1932, Monograph No. 15, National Survey of Secondary Education. Washington: Office of Education, 1933. Pp. vi + 72.

Zook, G. F., and others. "Report of the Committee on Revision of Standards," *North Central Association Quarterly,* VII (September, 1932), 185-205.